Empirical semi-groups and calibration

Vlad Bally, Emmanuel Temam

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Empirical semi-groups and calibration

Vlad Bally — Emmanuel Temam

N° 4873

Juillet 2003

THÈME 4

ISSN 0249-6399 ISRN INRIA/RR--4873--FR+ENG

Empirical semi-groups and calibration

Vlad Bally , Emmanuel Temam

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet MATHFI

Rapport de recherche n° 4873 — Juillet 2003 — 41 pages

Abstract: We present a probabilistic method to calibrate markovian model to a finite set of observed options prices. We prove that under a non-arbitrage condition, the underlying process is markovian. We describe an algorithm evolutive which allow us to compute the whole transition matrix of the underlying process from the option prices. Finally we discuss the efficiency of our algorithm.

Key-words: Markov process, inverse problem, model calibration, option pricing

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex (France)
Téléphone : +33 1 39 63 55 11 — Télécopie : +33 1 39 63 53 30

Semi-groupe empirique et Calibration

Résumé : Nous présentons une méthode probabiliste pour calibrer des modèles markoviens à des prix d'options observés. Nous établissons tout d'abord un cadre théorique en montrant qu'une condition de non-arbitrage spécifique permet de décrire le processus sous-jacent comme un processus de Markov. Ensuite, nous mettons au point un algorithme évolutif permettant de calculer les probabilités de transition infnitésimal du sous-jacent à partir des prix côtés sur le marché. Enfin, nous discutons de l'efficacité de l'algorithme

Mots-clés : Processus de Markov, problème inverse, calibration de modèle, évaluation d'option

Contents

List of Figures

1 Introduction

In the classical problem of calibration, the assets S_t^0 and S_t are given: S^0 is a bank account (with a known interest rate r) and S is a stock. Its dynamic follows the Black Scholes model of volatility σ. The aim is to find σ such that the call option prices computed with this volatility fit the best the empirical data - that is a finite number of call option prices which are traded on the market. But it is well known that this is not possible in practice. A call option of different maturities or different strikes give different implied volatilities - this is the well known volatility smile ([7, 8]). So the Black-Scholes model with constant volatility does not provide a satisfactory explanation of the prices observed on the market. Then a large variety of extensions of the standard Black Scholes model appeared, the aim of all these models being to overcome the above difficulty: local volatility model (Dupire's model), jump type diffusions, stochastic volatility diffusions and so on. All these models implied calibration technics based either on the partial differential equation (see [11, 10, 4] for Dupire models and [1, 12] for jump diffusion models) either on the minimization of relative entropy ([5, 2]).

A first question is about the legitimacy of a model or another - beyond the assertion that it explains well the empirical price table, because this is just the starting point of a long discussion about the significance of the 'well explaining'. Such a discussion is surely very interesting and maybe it represents the center of all the work concerning calibration. But our question is somehow more theoretical: how far are we allowed to push our imagination in order to produce market models which explain more or less well the empirical prices? What is the natural frame in which such a diversity of models live? A second subsequent question concerns completeness: jump models and stochastic volatility models are not complete. Is this a real difficulty or one may accept (or even privilege) this type of models? This problem has been already discussed in the interest rates theory. In this frame incompleteness is a structural characteristic of the market and one has not the choice between accepting or rejecting incomplete models. The answer which has been given there is the following: introducing the so called risk prime fixed by the market. This new free parameter absorbs the incompleteness. Put it otherwise: anyway the Black Scholes model is based on a free parameter, which is the volatility. Why not considering two free parameters instead of one? So, facing the problem of an infinite number of risk neutral probabilities, one parameterizes this family on a new parameter - the risk prime - and this parameter becomes itself subject to calibration. The common sense fact which is behind this solution is that, does not matter mathematical discussions, the market produces prices (at list in a fluid market) and the market price table is the basic object which has to be explained by the models.

So our starting point is the idea that the market price table produces automatically a "pricing machinery" which represents the natural underlying model. This leads to a non parametric point of view. Then different specific models appear as parametric approach to the problem - and parameterization is a way to reduce the number of unknowns in the calibration problem (which is naturally a sub-determined problem). We come back on this further on. Let us be more precise on the pricing machinery. We assume that the market is based on a bank account S_t^0 and on d stocks $S_t = (S_t^1, ..., S_t^d)$.Moreover we assume that

for every continuous positive function $\phi : R^d \to R_+$ and every $0 \leq t \leq T, x \in R^d$, a number $\Pi_{t,T}(\phi)(x)$ is given. This is the price of an European option of maturity T and payoff ϕ, at time t, if the stock price is $S_t = x$. So we dispose of a complete price table for the European options (which of course is not the case in practical situations - this is just a theoretical assumption). The function $x \to \Pi_{t,T}(\phi)(x)$ is assumed to be continuous so that $\Pi_{t,T}$ appears as an operator which maps continuous functions into continuous functions. Under the non arbitrage assumption we prove that this operator has to be linear, positive (and consequently monotone) and passes to monotone limits. Then Daniell's theorem asserts that we may represent it by a finite and positive kernel $\mu_{t,T}(x, dy)$, that is $\Pi_{t,T}(\phi)(x) = \int \phi(y)\mu_{t,T}(x, dy)$. This kernel has the martingale property $\int y\mu_{t,T}(x, dy) = x$. Moreover the above family of operators is a semi-group: they satisfy the Chapman Kolmogorov equation $\Pi_{t,T} = \Pi_{t,s} \circ \Pi_{s,T}$ for every $t \leq s \leq T$. We also prove that if the family of operators $\Pi_{t,T}$ has the above properties, then the market is arbitrage free. So these represent necessary and sufficient conditions for non arbitrage.The decision of working with a jump model or with a local volatility model does not appear as a natural consequence of arbitrage arguments but as a modeling hypothesis which may be (or not) benefit from computational reasons. Note that the stochastic volatility model does not enter in the frame described above: if $dS_t = \sigma(Y_t)S_t dW_t$ where Y is a diffusion process independent of W, then S is not a Markov process and so prices may not be given by a semi-group. This is because $\Pi_{t,T}(\phi)$ does not depend on S_t only but on Y_t as well. So markovianity enters in our hypothesis by mean of the assumption that the price of the option at time t is completely determined by the stock price at this moment.

Up to this moment no probability space appeared and one may wonder if this is really necessary or one may restrict himself to a deterministic calculus based on the above semi-group. The answer is beyond our competency but we think that even if this is possible, it does not seem desirable. Not only stochastic calculus technical permit to solve problems but probability theory is the natural language in which number of interesting financial questions are asked. So we go a step further and call up the Markov processes theory which, under mild regularity assumptions, provides a representation of the above semi-group as expectations of a Markov process. The probabilistic representation comes on naturally.. But we have to recall that the general Markov process theory represents a weak approach (in law) and does not automatically produce a unique probability space on which everything is going on, but describes the evolution of the underlying process by means of a whole family of probability measures on the canonical space of trajectories. So we are still far from a stochastic calculus (a substitute of stochastic calculus is available for symmetric Markov processes in Fuckushima's theory). We stop here with our axiomatic discussion and go on to calibration problems.

Since there is no underlying model, searching for the volatility is meaningless. The object to be determined in our frame is the family of positive measures $\mu_{t,T}(x, dy)$. We have the

following structural information on them:

$$\begin{aligned}(P)\quad \mu_{t,T}(x,D) &= e^{-r(T-t)}, \quad \mu_{t,T}(x,dy)\geq 0,\\ (M)\quad \int y\mu_{t,T}(x,dy) &= x,\\ (S)\quad \int \phi(y)\mu_{t,T}(x,dy) &= \int\int \phi(z)\mu_{s,T}(y,dz)\mu_{s,t}(x,dy).\end{aligned}$$

(P) means that up to a normalization $\mu_{t,T}(x,dy)$ is a probability measure, (M) is the martingale property and (S) is the semi-group property. Except for this we have some experimental prices, that means a table of prices $C_{0,T_k}(x_0,K_l), k=1,...,n, l=1,...,m$. These are call option prices of maturity T_k and strike K_l. In our numerical experiments we take $n=4$ and $m=5,10,20$. If we want that our measures explain this prices we have the equalities:

$$(C_{k,l})\quad C_{0,T_k}(x_0,K_l)=\int (y-K_l)_+\mu_{0,T_k}(x_0,dy).$$

The problem of finding the family of measures which verifies these properties is an infinite dimensional non parametric problem and it is obvious that this problem is dramatically sub determined. So the problem is now to find a way to reduce in a reasonable way the number of degrees of freedom. One way of doing this is to suppose that the dynamics of the underling stock is given by some model (Dupire's model, jump type diffusions....) and then one comes on in a parametric frame: the parameters would be the local volatility or the jump measure for example. Even in this case the space of parameters is still infinite dimensional. Thus we have to consider a discretization procedure in order to come to a numeric problem. But, a model is a way of including in our computations some restrictions coming from some a priori assumptions on the dynamics of the stock and so one search for the solution in a smaller space.. This point of view has the advantage that the restrictions produced by the model have a clear intuitive meaning, so we know what we are doing. But what is much less clear is that what we are doing is what we have to do - because we do not know that the empirical prices come from a specific model, namely the one we have chosen.. So our point of view is to avoid model hypothesis and to postpone a priori restrictions as long as possible. We work directly with a discretization of the family of measures $\mu_{t,T}$ without any model hypothesis.

The first discretization concerns time: we consider a time grid $0=t_0<...<t_n=T$, typically $n=12$. We assume that the empirical call prices are known for all these epochs. We mentioned before that only four epochs are given but we extent the data to twelve epochs just by linear interpolation - and numerical experiments show that this works very well and does not represent a real problem. Then we use the semi-group hypothesis in order to write the restriction $(C_{k,l})$ under the form

$$\begin{aligned}(C_{k+1,l})\quad C_{0,t_{k+1}}(x_0,K_l) &= \int (y-K_l)_+\mu_{0,t_{k+1}}(x_0,dy)\\ &= \int (y-K_l)_+\mu_{t_k,t_{k+1}}(z,dy)\mu_{0,t_k}(x_0,dz).\end{aligned}$$

Our algorithm is evaluative. At the step $k+1$ we assume that we have already computed μ_{0,t_k} at step k and we want to compute $\mu_{t_k,t_{k+1}}$. So we consider $(C_{k+1,l}), l = 1, ...m$ as a system of equations with the unknown $\mu_{t_k,t_{k+1}}$. Once $\mu_{t_k,t_{k+1}}$ computed, we use the Chapman Kolmogorov equation in order to produce $\mu_{0,t_{k+1}} = \mu_{0,t_k} \otimes \mu_{t_k,t_{k+1}}$ and go further to the next step.

Recall that except for the above equations we also have the conditions given by (P), and (M). We are still in an infinite dimensional setting and we have to perform one more discretization. We consider a space grid $0 < y_1 < ... < y_M$ and replace $\mu_{t_k,t_{k+1}}(y_i, dy)$ by an approximate $\sum_{j=1}^{M} \pi^{ij}_{k,k+1} \delta_{y_j}(dy)$. Now our equations read

$$
\begin{aligned}
(P_i) \quad \sum_{j=1}^{M} \pi^{ij}_{k,k+1} &= e^{-r(t_{k+1}-t_k)}, i = 1, ..., M \quad \pi^{ij}_{k,k+1} \geq 0, \\
(M_i) \quad \sum_{j=1}^{M} y_j \pi^{ij}_{k,k+1} &= y_i, i = 1, ..., M \\
(C_{k+1,l}) \quad C_{0,t_{k+1}}(x_0, K_l) &= \sum_{j=1}^{M} (y_j - K_l)_+ \pi^{ij}_{0,k+1} \\
&= \sum_{j=1}^{M} (y_j - K_l)_+ \sum_{p=1}^{M} \pi^{0,p}_{0,k} \pi^{p,j}_{k,k+1}, l = 1, ..., m.
\end{aligned}
$$

We have now $2M + m$ equations with $M \times M$ unknowns and so we have still to reduce the number of degrees of freedom.. One idea is to use a three branches tree, but from a numerical point of view it seems not the best possible idea ([6]). The reason is that such a tree is extremely sensible to the location of the points $y_i, i = 1, ..., M$. In order to smooth our algorithm we instead use a finite element type method. More precisely we fix i and look to $\pi^{ij}_{k,k+1}$ as a function in the forward argument that is $\pi^{ij}_{k,k+1} = \pi^{i}_{k,k+1}(y_j)$ and then project this function on three trials $\psi_{i,p}, p = 1, 2, 3$. So we have $\pi^{i}_{k,k+1}(y_j) = \sum_{p=1}^{3} \lambda^i_p \psi_{i,p}(y_j)$. Now the unknowns are $\lambda^i_p, i = 1, ..., M, p = 1, ..., 3$. We have now a system of $2M + m$ equations with $3M$ unknowns. In our numerical tests we took $M = 150$ and $m = 5, 10, 20$ so this system is slightly undetermined. One way to countering this difficulty is to interpolate and so to obtain all the prices $C_{0,t_{k+1}}(x_0, y_j), j = 1, ..., M$ and then to solve a system of $3M$ equations with $3M$ unknowns. Numerical experiments show that the interpolation works very well and the system is well solved. But the results are bad.

Let us explain a little bit what we mean by numerical experiments. We considered synthetic data produced by a local volatility (Dupire) diffusion with four types of volatilities $\sigma = 0.3, \sigma(t, x) = 15/x, \sigma(t, x) = 0.05 + 0.1 \exp(-x/100) + 0.01t$ and $\sigma(t, x) = 0.2\mathbf{1}_{x \notin [90,110]} + 0.4\mathbf{1}_{x \in [90,110]}$. In each case we produced the corresponding call prices $C_{0,t_{k+1}}(x_0, K_l)$ and solved the above system of linear equations. Then we asked the question if the results are good or not. There are several ways to appreciate the quality of the results. The first one is to see if we are able to feet well the experimental (synthetic) data, that is to satisfy the

equations $(C_{k,l})$. And this works well. The second way, which is more subtle, is to produce by our algorithm the weights $\pi^{ij}_{k,k+1}$ and then to use these weights in order to compute put option prices. And then to compare our results with the prices given by the synthetic underlying model. And this works well also. The third test consists in computing the local volatility σ^i_k associated to the weights $\pi^{ij}_{k,k+1}, j = 1, ..., M$ (the quadratic variation) and then to compare it with the theoretical one $\sigma(t_k, y_i)$. And this works bad. The reason is the following. One may produce two different local volatilities σ' and σ'' such that the corresponding call prices $C'_{0,t_{k+1}}(x_0, K_l)$ and $C''_{0,t_{k+1}}(x_0, K_l)$ are different but very close each another. As a consequence although the system of equations is well solved (we succeed to feet very well the experimental data), we are not able to distinguish between σ' and σ''. This seems natural: as long as we just feet prices we are at the precision level of prices and we may hope that our results are sufficiently accurate in order to compute other prices. And this is true. But if we go deeper and want to find the underlying volatility then we need another level of accuracy. In fact, if the volatility is well computed, then we may compute Greeks, and these are no more prices but derivatives of prices which are much more sensible quantities. We conclude that if we want to achieve a precision level which is compatible with volatility, fitting prices is not a sufficient criterion and we have to employ another criterion which is compatible with this level of accuracy. This is why we turn to another setting in which we do no more compare prices but implied volatilities, and this amounts to solve the following non linear minimization problem.

Recall that our unknowns are $\lambda^i_p, i = 1, ..., M, p = 1, 2, 3$. We denote $\lambda = (\lambda^1, ..., \lambda^M) = (\lambda^1_2, ..., \lambda^M_2)$, that is the weights of the central trials $\psi_{2,i}, i = 1, ..., M$. Then we solve explicitly the equations (P_i) and (M_i) and so we obtain $\lambda^i_1 = \lambda^i_1(\lambda)$ and $\lambda^i_3 = \lambda^i_3(\lambda)$. So our unknown is now $\lambda = (\lambda^1, ..., \lambda^M)$. To each such a λ we associate the weights $\pi^{ij}_{k,k+1}(\lambda)$ and then the call option prices given by these weights, that is $C_k(\lambda) = \sum_{j=1}^M (y_j - K_l)_+ \pi^{ij}_{0,k}(\lambda)$. Finally we denote by $Iv_{k,l}(\lambda)$ the implied volatility of $C_k(\lambda)$. Then we consider $\overline{Iv}_{k,l}$ to be the implied volatility of the corresponding experimental call option price $C_{0,t_{k+1}}(x_0, K_l)$. Finally we consider the coast function

$$c(\lambda) = \sum_{l=1}^m \left| Iv_{k,l}(\lambda) - \overline{Iv}_{k,l} \right|^2 + p \sum_{i=1}^M (\left| \lambda^i - \lambda^i_1(\lambda) \right|^2 + \left| \lambda^i - \lambda^i_3(\lambda) \right|^2)$$

where p is a positive weight. The important part of our coast function is that containing the implied volatility. The second term has somehow a regularisation effect - it contributes to a uniform distribution of the weights on the three trials. But one may conceive another type of regularisation term as well. Then we use a quasi Newton algorithm in order to solve the following problem: find $\lambda^* = \arg\min c(\lambda)$ under the constraint $\pi^{ij}_{k,k+1}(\lambda) \geq 0$. And this gives good results.

We close our introduction with some commentaries.. We claimed in the beginning that our algorithm is model free. And it is. The fact that we employ the implied volatility is not a model hypothesis but a scale which is used in order to interpret data. There is another point in which an a priori guess on the dynamics of the underlying stock comes on: in the

choice of the grid $y_i, i = 1, ..., M$ and of the trials $\psi_{i,p}, i = 1, ..., M, p = 1, 2, 3$. We chose an "exponential grid" because we expect an exponential behavior of the stock. And we use the implied volatility associated to the empirical prices in order to construct the trials. So the geometry of our grid and of our trials supposes a "Black-Scholes type" behavior of the stock. This may be seen as a light counterpart of model hypothesis and has the advantage of being very flexible.

2 European options price tables

We assume that on the market is given a bank account S_t^0 and d stocks $S_t = (S_t^1, ..., S_t^d)$. S^0 evolutes in a deterministic way according to $S_t^0 = S_0^0 e^{rt}$ where r is the deterministic interest rate. S represents a risky stocks and we know nothing about their evolution. Usually one assumes that it evolutes according to some stochastic equation so there is some probabilistic model which gives the behavior of S_t. The uncertainty related to S is expressed by means of this probabilistic model. But here we try to see what can be said without any underling model.

We consider an open set $D \subseteq R^d$ (typically $D = \{(x^1, ..., x^d) : x^i > 0, i = 1, ..., d\}$) and denote by C_+ the space of continuous functions defined on D and taking values in $R_+ := (0, \infty)$. Any continuous function $\phi \in C_+$ is thought to be a payoff. A $(T, \phi)-$option is a contract which gives the right to the owner to a payment of $\phi(S_T)$ (exactly) at time T. Our assumptions on the market are the following:

Assumption (H_1): The market is completely "fluid" in the sense that

i) At any time $t \geq 0$ one may buy and sell any quantity of stock S at price S_t (this price is not known before t, but is known at time t). One also may borrow or lend at time t any given quantity of stock S for a given period $T > t$.

ii) One may borrow or lend money to the bank (any quantity) with interest rate r (both for borrowing or lending).

iii) One may sell or buy at any moment $0 \leq t \leq T$ any $(T, \phi)-$ option for any payoff function $\phi \in C_+$.

We consider now a price operator $\Pi_{t,T}(\phi)(x)$ which represents the price of a $(T, \phi)-$ option at time $t \leq T$, if $S_t = x$. Our hypothesis (H_1, iii) implies that $\Pi_{t,T}(\phi)(x)$ is known for $x = S_t$ only but we think that S_t may take any positive value and so we will assume that $\Pi_{t,T}(\phi)(x)$ is known for every $x \in (0, \infty)$.

We define a price machine Π to be a family of operators $\Pi_{t,T} : C_+ \to C_+$ for every $0 \leq t \leq T < \infty$.

We will now define an arbitrage opportunity (free lunch). We say that Π admits an arbitrage opportunity if there exists $0 \leq t \leq T$, $\theta \in D$ and $\varepsilon > 0$ such that an agent may buy and sell the stock $S_s, s \in [0, T]$, borrow and lend money at the bank and buy and sell $(T, \phi)-$ options at any moment $s \in [0, T]$, for any payoff $\phi \in C_+$ at price $\Pi_{t,T}(\phi)(S_t)$ in such a way that:

(H_1) For every possible evolution of S_t, the agent does not loose money.

$(\mathbf{H_2})$ If $|S_t - \theta| \leq \varepsilon$ then the agent wins a strictly positive amount of money..

In other words: all the treading operations presented in $H_1, i), ii), iii)$ are allowed and the price of a $(T, \phi)-$option at time $s \in [0, T]$ is $\Pi_{t,T}(\phi)(S_s)$. An arbitrage opportunity means that an agent may trade in such a way that he is sure that he does not lose money (property (H_1)) and, in some "favorable situation" - described by the fact that the price of the stock is closed to a given value θ - he wins money. Note that we do not describe this opportunity by "$S_t = \theta$" but we just ask the price S_t to be close to θ up to some $\varepsilon > 0$. This way of taking things is motivated by the common sense assertion that we may not expect that the event $S_t = \theta$ really occurs while we may hope, with "non null probability" that the event $|S_t - \theta| \leq \varepsilon$ occurs.. So the arbitrage opportunity is effective with strictly positive probability.. As we mentioned in the beginning no probability space is given so this is just a probabilistic intuition which motivates the definition. So the definition of the arbitrage opportunity obliges us to assume that $x \to \Pi_{t,T}(\phi)(x)$ is continuous.

Let us give a more quantitative description of an arbitrage. An operation done at time t and sold at time T is described by the following objects. First of all one considers a number $\alpha \in R$ which represents the quantity of stock S_t which is traded. If $\alpha < 0$ this means that the agent buys a quantity $-\alpha$ of stock - after this operation he has $\alpha S_t < 0$ dollars. If $\alpha > 0$ then this means that the agent sails a quantity α of stock, and after this he has $\alpha S_t > 0$ dollars. If the agent buys, at time T he will sail the same quantity of stock at time T and so he will get $-\alpha S_T$. The situation is a little bit different if the agent sails a quantity of stock - for the simple reason that he owns no stock.. So, in order to sail a quantity α of stock he has to borrow it, and then, at time T he has to pay αS_T in order to honor his duty.. Next one considers the payoffs $\phi_i, i = 1, ..., n$ and the numbers $\beta_i, i = 1, ..., n$. The agent trades the $(T, \phi_i)-$ options and β_i represents the quantity of option which is traded. As before, if $\beta_i > 0$ this means that the agent sails β_i $(T, \phi_i)-$ options and so he wins $\beta_i \Pi_{t,T}(\phi_i)(S_t) > 0$ dollars. At moment T he has to honor these options and so he pays $\beta_i \phi_i(S_T)$. If $\beta_i < 0$ this means that the agent buys a quantity $-\beta_i$ of $(T, \phi_i)-$ options and so his gain is $\beta_i \Pi_{t,T}(\phi_i)(S_t) < 0$ dollars. At time T he wins $-\beta_i \phi_i(S_T) > 0$. We introduce now two functions

$$\begin{aligned} g_t(S_t) &= \alpha S_t + \sum_{i=1}^{n} \beta_i \Pi_{t,T}(\phi_i)(S_t), \\ s_T(S_T) &= -\alpha S_T - \sum_{i=1}^{n} \beta_i \phi_i(S_T). \end{aligned}$$

Having in mind the above discussion, $g_t(S_t)$ is the gain of the agent at time t (the moment when the operation starts) and $s_t(S_T)$ represents the sold of the operation at final time T and the agent has to pay this sum. We say that this operation represents an arbitrage opportunity if there is some $\theta \in D$ and $\varepsilon > 0$ such that $g(x) > 0$ for $x \in B_\varepsilon(\theta)$ and $s_T(x) \leq 0$ for every $x \in R$. If such an opportunity exists, then one may achieve an arbitrage in the following way: up to time t he does nothing ant at time t he checks if $S_t \in B_\varepsilon(\theta)$. If this is not the case he does nothing, but if this is true, then he buys/sails α stocks and β_i $(T, \phi_i)-$ options. His gain is $g_t(S_t) > 0$. At time T he has to pay $s_T(S_T) \leq 0$ so his gain is larger or equal to $g_t(S_t) > 0$.

Finally we give a technical hypothesis which essentially says that prices are not expected to be larger then a sufficiently large level. This is an analogues of the thighteness property for measures.

(**H$_3$**) For every $\delta > 0$ and $t \in [0,T)$ there exists some $K_{t,\delta} > 0$ such that for every $\phi \in C_+$, with $0 \leq \phi \leq 1$ and such that the support of ϕ is included in $B^c_{2K_{t,\delta}}(0)$, one has $\Pi_{t,T}(\phi)(x) \leq \delta$ for every x such that $|x| \leq K_{t,\delta}$.

Here $B_r(x) := \{y : |x-y| \leq r\}$. The probabilistic interpretation of the above hypothesis is that for any $\delta > 0$ there exists some K such that the probability that $|S_t - S_0| \geq K$ is smaller then K. If we express this by means of some continuous functions ϕ (which gives a more complicated statemanet) this is because in the beginning we decided to work with continuous functions. So ϕ has to be seen as the regularization of the indicator function $1_{[K,\infty)}$.

Lemma 2.1 *Suppose that* Π *does not admit arbitrage opportunities.. Then for every* $\alpha \geq 0, \phi, \psi \in C_+$

$$\begin{aligned} a) \quad \Pi_{t,T}(\phi+\psi)(x) &= \Pi_{t,T}(\phi)(x) + \Pi_{t,T}(\psi)(x) \\ b) \quad \Pi_{t,T}(\alpha\phi)(x) &= \alpha\Pi_{t,T}(\phi)(x) \\ c) \quad \phi &\leq \psi \Rightarrow \Pi_{t,T}(\phi) \leq \Pi_{t,T}(\psi), \\ d) \quad \phi_n &\downarrow \phi \Rightarrow \Pi_{t,T}(\phi_n) \downarrow \Pi_{t,T}(\phi), \\ e) \quad \Pi_{t,T}(1)(x) &= e^{-r(T-t)}. \end{aligned}$$

The property d*) holds true for every sequence* $\phi_n, n \in N$ *such that* ϕ_1 *is bounded.*

Proof. a) Suppose that $\Pi_{t,T}(\phi+\psi)(\theta) > \Pi_{t,T}(\phi)(\theta) + \Pi_{t,T}(\psi)(\theta)$ for some given θ. Since the above functions are continuous one may find some $\varepsilon > 0$ such that the inequality holds true for every $x \in B_\varepsilon(\theta)$. Then one trades in the following way. Up to t one does nothing ant at time t one checks if $S_t \in B_\varepsilon(\theta)$. If not, he does nothing. If yes, then he seals a $(T, \phi+\psi)-$option and buys an $(T,\phi)-$option and a $(T,\psi)-$option. The sold of these operations is $g_t(S_t) = \Pi_{t,T}(\phi+\psi)(S_t) - (\Pi_{t,T}(\phi)(S_t) + \Pi_{t,T}(\psi)(S_t)) > 0$. He keep this gain. At time T he receives $(\phi+\psi)(S_T)$ because he exercises his $(T,\phi+\psi)-$option and he gives this money because he has to honor the two options that he had sold. His gain is $g_t(S_t) > 0$. So we have an arbitrage opportunity.. The proof of b) is similar.

Let us prove c). Suppose that $\Pi_{t,T}(\phi)(\theta) > \Pi_{t,T}(\psi)(\theta)$, and consequently this holds true on a whole $B_\varepsilon(\theta)$. Suppose also that $S_t \in B_\varepsilon(\theta)$. Then we sell a $(T,\phi)-$option and by an $(T,\psi)-$option in order to obtain $g_t(S_t) = \Pi_{t,T}(\phi)(S_t) - \Pi_{t,T}(\psi)(S_t) > 0$. At time T we have to pay $s_T(S_T) = \phi(S_T) - \psi(S_T) \leq 0$.

We prove d). Recall that t is fixed. Suppose that there is some θ such that $\inf_n \phi_n(\theta) > \delta + \phi(\theta)$ for some $\delta > 0$. Suppose also that $M \geq \phi_1 \geq 0$. Then we take $\delta' = \delta/2M$ and consider $K_{t,\delta'}$ from the hypothesis (H_3). Since $\phi_n \downarrow \phi$ we may use Dini's theorem and conclude that the convergence is uniform on compact sets. So we may find n_δ such that

$$0 \leq \phi_{n_\delta}(x) - \phi(x) \leq \frac{\delta}{4} \quad for \quad |x| \leq K_{t,\delta'} + 1.$$

We will trade on the payoff ϕ_{n_δ} which is now fixed. We find some $\varepsilon > 0$ such that $\Pi_{t,T}(\phi_{n_\delta})(x) > \delta + \Pi_{t,T}(\phi)(x)$ for $x \in B_\varepsilon(\theta)$. We also consider a localization function $\chi \in C$ such that $0 \leq \chi \leq 1, \chi(x) = 0$ for $|x| \leq K_{t,\delta'}$ and $\chi(x) = 1$ for $|x| \geq K_{t,\delta'} + 1$. Then we construct $\psi = \chi(\phi_{n_\delta} - \phi)$. Note that, since $0 \leq \phi_{n_\delta} - \phi \leq \phi_1 \leq M$, we have

$$\Pi_{t,T}(\psi)(x) \leq \Pi_{t,T}(M\chi)(x) = M\Pi_{t,T}(\chi)(x) \leq M \times \frac{\delta}{2M} = \frac{\delta}{2}.$$

We are now ready to trade. We are at time t and suppose that $S_t \in B_\varepsilon(\theta)$. Then we buy a (T, ϕ)-option and a (T, ψ)-option and sell a (T, ϕ_{n_δ})-option. The gain is

$$g_t(S_t) = \Pi_{t,T}(\phi_{n_\delta})(S_t) - \Pi_{t,T}(\phi)(S_t) - \Pi_{t,T}(\psi)(S_t) > \delta - \frac{\delta}{2} = \frac{\delta}{2}.$$

At time T we have to pay $s_T(S_T) = \phi(S_T) + \psi(S_T) - \phi_{n_\delta}(S_T)$. If $S_T \leq K_{t,\delta'} + 1$ then $|\phi(S_T) - \phi_{n_\delta}(S_T)| \leq \frac{\delta}{4}$ and $\psi(S_T) = 0$ so $g_t(S_t) - s_T(S_T) > \frac{\delta}{2} - \frac{\delta}{4} = \frac{\delta}{4}$. If $S_T > K_{t,\delta'} + 1$ then $\chi(S_T) = 1$ and so $\phi(S_T) + \psi(S_T) - \phi_{n_\delta}(S_T) = \phi(S_T) - \phi_{n_\delta}(S_T) + \chi(S_T)(\phi(S_T) - \phi_{n_\delta}(S_T)) = 0$. So in both situations we have a strictly positive gain.

The proof of $e)$ is trivial, trading on the bank account.. So the proof is completed. □

We are now ready to produce a measure which represents $\Pi_{t,T}$. First of all we extend this operator from positive functions to real functions in the following standard way. If $f = g - h$ for some positive functions g and h then we define $\Pi_{t,T}(f) := \Pi_{t,T}(g) - \Pi_{t,T}(h)$. Note that this definition does not depend on the decomposition of f in $g - h$. In fact, if $f = g' - h'$ then we claim that $\Pi_{t,T}(g) - \Pi_{t,T}(h) = \Pi_{t,T}(g') - \Pi_{t,T}(h')$. Having in mind the linearity on positive functions this amounts to $\Pi_{t,T}(g + h') = \Pi_{t,T}(g' + h)$. But $g + h' = g' + h$ and so this is true. So we have a correct definition of our operator on the whole space of the continuous functions and the extended operator inherits the properties of our initial operator: it is linear, monotone and pass to decreasing limits. Then Daniell's theorem (see [R]) says that, for each fixed x, the functional $\phi \to \Pi_{t,T}(\phi)(x)$ may be represented by a positive Radon measure $\mu_{t,T}(x, dy)$ of total mass e^{-rt}. We have one more problem about the measurability of $x \to \mu_{t,T}(x, A)$ where $A \subseteq R$ is some measurable set (this property enters in the definition of a transition kernel and is necessary in order to have semi-group properties, as we will discuss in a moment). The proof of this property is standard. One denotes by $\mu_{t,T}(x, \phi)$ the integral of ϕ with respect to $\mu_{t,T}(x, dy)$. If ϕ is continuous then $\mu_{t,T}(x, \phi) = \Pi_{t,T}(\phi)(x)$ is continuous and so it is also measurable. Consider now a closed set A. Then there exists a sequence of continuous functions ϕ_n such that $\mu_{t,T}(x, \phi_n) \downarrow \mu_{t,T}(x, A)$ for every x. This is a regularity property for Radon measures and Daniell's theorem produces a Radon measure. So $x \to \mu_{t,T}(x, A)$ is measurable for closed sets. Finally the measurability property follows for general Borel sets by a monotone class argument. So we have proven:

Proposition 2.1 *If there is no arbitrage opportunity then there exists a positive kernel* $\mu_{t,T}(x, dy)$ *such that* $\mu_{t,T}(x, R^d) = e^{-rt}$ *and*

$$\Pi_{t,T}(\phi)(x) = \int \phi(y)\mu_{t,T}(x, dy).$$

We give now the semi-group property and the martingale property.

Lemma 2.2 *Suppose that there is no arbitrage opportunity.. Then*

$$\begin{aligned} f) \quad \Pi_{t,T}(I)(x) &= x \quad where \quad I(y) = y \\ g) \quad \Pi_{s,t}(\Pi_{t,T}(\phi))(x) &= \Pi_{s,T}(\phi)(x) \quad \forall s < t < T. \end{aligned}$$

Proof. Suppose that $\Pi_{t,T}(I)(\theta) > \theta$ for some θ. Then the inequality holds true on a whole ball $B_\varepsilon(\theta)$. If $S_t \in B_\varepsilon(\theta)$ then we sell a $(T, I)-$option and buy a stock. The gain is $g_t(S_t) = \Pi_{t,T}(I)(S_t) - S_t > 0$. At time T we sell the stock and receive S_T.We give $I(S_T) = S_T$ to the owner of the option and so our global gain is strictly positive. If $\Pi_{t,T}(I)(\theta) < \theta$ for some θ and $S_t \in B_\varepsilon(\theta)$, we buy an $(T, I)-$option and borrow a unity of stock which we sail and we obtain the sold $g_t(S_t) = S_t - \Pi_{t,T}(I)(S_t) > 0$. At time T we exercise our option and obtain a unity of stock that we give to the person from which we have borrowed..

Let us prove $g)$. Suppose that $\Pi_{s,t}(\Pi_{t,T}(\phi))(\theta) > \Pi_{s,T}(\phi)(\theta)$. Once again the inequality holds true on a whole ball $B_\varepsilon(\theta)$. One does nothing up to s and if $S_s \in B_\varepsilon(\theta)$ then he seals a an $(t, \psi)-$option where $\psi = \Pi_{t,T}(\phi)$. He buys then an $(T, \phi)-$option. The sold of these operations is $\Pi_{s,t}(\Pi_{t,T}(\phi))(S_s) - \Pi_{s,T}(\phi)(S_s) > 0$. At time t he sells his $(T, \phi)-$option and receives $\Pi_{t,T}(\phi)(S_t)$ which is exactly the sum he has to pay to the owner of the $(t, \psi)-$obtain that he has sold. □

We are now able to give our main result.

Theorem 2.1 *Given the family of numbers $\prod_{t,T}(\phi)(x), 0 \leq t \leq T, \phi \in C, x \in R$ the following two assertions are equivalent:*

A. There exists a family of positive kernels $\mu_{t,T}(x, dy), 0 \leq t \leq T$, such that the price operators are expressed as

$$\Pi_{t,T}(\phi)(x) = \int \phi(y)\mu_{t,T}(x, dy)$$

and these measures verify the martingale condition $f)$.

B. The family of price operators $\Pi_{t,T}$ admit no simple arbitrage.

ii) Suppose that the above assertions hold true. Then $\Pi_{t,T}, 0 \leq t \leq s \leq T$ satisfies the Chapman Kolmogorov equation $\Pi_{t,T} = \Pi_{t,s} \circ \Pi_{s,T}$.

Proof. i) We have already shown that $B \Rightarrow A$. Let us now suppose that A holds true. Then, using first the martingale property and then the representation by means of the positive measure $\mu_{t,T}$ we obtain

$$\begin{aligned} g_t(x) &= \alpha x + \sum_{i=1}^{n} \beta_i \Pi_{t,T}(\phi_i)(x) = \alpha \Pi_{t,T}(I)(x) + \sum_{i=1}^{n} \beta_i \Pi_{t,T}(\phi_i)(x) \\ &= \int \mu_{t,T}(x, dy)(\alpha y + \sum_{i=1}^{n} \beta_i \phi_i(y)) = -\int \mu_{t,T}(x, dy) s_T(y) \leq 0. \end{aligned}$$

This does not permit to obtain an arbitrage opportunity.. □

3 Probabilistic representation

In the previous section no probability representation was supposed. But it is well known that a semi-group as the one presented there is always the transition semi-group of a Markov process, so that a natural probabilistic interpretation comes on. A first way of taking things is to give an initial law μ_0 and then to construct a stochastic process which has μ_0 as initial law and $\Pi_{t,T}$ as transition semi-group in the following way. Let $\Omega = \{\omega : [0,\infty) \to R^d : t \to \omega(t)$ is right continuous and has left hand limits$\}$ be the canonical space of trajectories and defines the coordinates process $X_t(\omega) = \omega(t)$ and the corresponding filtration $F_t = \sigma(X_s : s \leq t\}$. Then the probability measure P^{μ_0} - under which $(X_t)_{t\geq 0}$ is a non-homogeneous Markov process - is obtained in the following way. One defines the cylindrical probabilities

$$\begin{aligned} P_{(t_1,...,t_n)}(X_0 &\in A_0, X_{t_1} \in A_1, ..., X_{t_n} \in A_n\} \\ &: = \int_{A_0} \mu_0(dx_0) \int_{A_1} \mu_{0,t_1}(x_0, dx_1) ... \int_{A_n} \mu_{t_{n-1}}(x_{n-1} dx_n) \end{aligned}$$

where $0 < t_1 < ... < t_n$ and $A_0, ..., A_n$ are Borel sets in R^d. Then one employs Kolmogorov's theorem in order to construct a probability measure P^{μ_0} on (Ω, F_∞) such that $P_{(t_1,...,t_n)}$ represents the law of $(X_0, X_{t_1}, ..., X_{t_n})$ under P^{μ_0}. It turns out that X is a (non-homogeneous) Markov process under this probability and $E^{\mu_0}(f(X_t) \mid F_s) = \Pi_{s,t}(f)(X_s)$, which means that $\Pi_{s,t}$ is the transition semi-group of X.

This procedure works in all generality but in order to obtain a nice theory one has to restrict himself a little bit. First of all one assumes homogeneity, that is

$$(\mathbf{H_4}) \quad \Pi_{t,T} = \Pi_{0,T-t}$$

which means that the price of an option depends on the time to maturity only. One also has to make the (rather natural) regularity assumption:

$$(\mathbf{H_5} \quad \lim_{t\to 0} \Pi_{0,t} f(x) = f(x)$$

for every $x \in R^d$ and every $f \in C_0$ where C_0 is the space of continuous real functions on R^d which vanished at infinity.

It is easy to check (as a consequence of the hypothesis (H_3)) that if $f \in C_0$ then $\Pi_{0,t} f \in C_0$ so we have a homogeneous semi-group of operators from C_0 to C_0 which satisfy the regularity condition (H_5). Then it is proved in [3] theorem (9.4) that the Markov process associated to such a semi-group is a standard Markov process and there exists a classical and well developed theory for this type of processes. We do not enter in more details but send the reader to [3] or [13] for a complete exposition of this theory. We just make some commentaries.. First of all, a standard process is described by a family of probability measures $P^x, x \in R^d$ on (Ω, F_∞) so that P^x represents the law of $(X_t)_{t\geq 0}$, if $X_0 = x$. In particular $dP^{\mu_0} = dP^x \mu_0(dx)$. This is a weak approach in the following sense. The classical model of Black and Scholes describes the evolution of the stock S (which in our context is

the Markov process X) by means of the stochastic equation $dS_t = \sigma S_t dW_t, S_0 = s_0$ where W is a Brownian motion. So one may define on the same probability space (Ω, F_∞, P) (the one where W is defined) all the processes $S_t^{s_0}, s_0 \in R^d$ and then P^{s_0} is the law of S^{s_0} under P. This means that for every $\omega \in \Omega$ we have an application $(t, s_0) \to S_t^{s_0}(\omega)$ which, under sufficient regularity assumptions on the coefficients of the underlying stochastic equation, is a flow of dipheomorfisms. This is a very strong and useful property (see [K] for a complete presentation of the theory related to stochastic flows). But a standard Markov process can generally not be represented by means of an underlying flow, and this is why we say that this is a weak approach. In particular there is no Brownian motion and no stochastic equation coming on. And so there is no stochastic Ito calculus. In the case of symmetric Markov processes (this means that $\int m(dx)\phi(x)\Pi_{0,t}(\psi)(x) = \int m(dx)\psi(x)\Pi_{0,t}(\phi)(x)$ for some measure m) a substitute of the stochastic calculus is settled (see [9]). But we do not discuss this here.

3.1 Call option price tables

In calibration problems we have not the whole price table but a few prices of call option prices. This motivates the following question. We are given a call price table and we want to use it in order to find the price machinery which produces this call prices. So we denote $C_{t,T}(x, K)$ the price of a Call option of maturity T and strike K at time t if the value of the underlying stock is $S_t = x$. In our previous notation we have $C_{t,T}(x, K) = \Pi_{t,T}(\theta_K)(x)$ where $\theta_K(y) = (y - K)^+$. Since the linear combinations of the functions of type θ_K are dense in the class of the differentiable functions it is clear that knowing the call options prices completely determine all the European option prices. But we would like to give a more precise result concerning approximation. For simplicity we restrict ourself to the one dimensional case that is $d = 1$. So we consider some $h > 0$ and the grid $x_k = kh, k \in N$ and want to approximate $\Pi_{t,T}(\phi)(x)$ by a linear combination of $C_{t,T}(x, x_k), k \in N$. We also prove that if $K \to C_{t,T}(x, K)$ is twice differentiable then $\mu_{t,T}(x, dy)$ is absolutely continuous with respect to the Lebesgue measure. We restrict ourself in this section to the one dimensional case - but it is clear that the multi-dimensional case may be treated in an analogues way.

Proposition 3.1 *i) Suppose that $C_{t,T}(x, K)$ is known for every $0 \leq t \leq T, x \in R_+, K \in R_+$. Then there exists a unique Π such that $C_{t,T}(x, K) = \Pi_{t,T}(\theta_K)(x)$. More precisely we may approximate the price $\Pi_{t,T}(\phi)(x)$ of any $(T, \phi)-$option of a differentiable payoff ϕ in the following way:*

$$i) \quad \Pi_{t,T}(\phi)(x) = \frac{1}{h}\sum_{k=0}^{\infty} \phi(x_k)(C_{t,T}(x, x_{k+1}) + C_{t,T}(x, x_{k-1}) - 2C_{t,T}(x, x_k)) + o(h)$$

where $|o(h)| \leq e^{r(T-t)} \|\phi'\|_\infty h$.

ii) For every twice differentiable positive payoff function ϕ one has

$$ii) \quad \Pi_{t,T}(\phi)(x) = \int_0^\infty \phi''(y) C_{t,T}(x, y) dy.$$

In particular, if the function $K \to C_{t,T}(x,K)$ is twice differentiable and has continuous second order derivatives then $\mu_{t,T}(x,dy) = \partial_y^2 C_{t,T}(x,y)dy$.

Remark 3.1 *As it is clear from the following proof the above properties have nothing to do with the dynamics of the stock underlying the call option prices but is a basic fact of distribution theory. The point is that the call option prices correspond to the special function $(x-K)_+$ and the second derivative of this function with respect to K is the Dirac mass in K.*

Proof. Let us denote ϕ_h the polygonal line approximation of ϕ defined by $\phi_h(x_k) = \phi(x_k)$ and ϕ_h piecewise linear. Consider the trials ψ_k defined by $\psi_k(x_k) = 1, \psi_k(x_{k-1}) = \psi_k(x_{k+1}) = 0$, ψ_k is zero outside $[x_{k-1}, x_{k+1}]$ and piecewise linear on this interval. Then it is easy to check that $\phi_h(x) = \sum_{k=0}^{\infty} \psi_k(x)\phi(x_k)$. Note also that $\psi_k(x) = \frac{1}{h}(\theta_{x_{k+1}} + \theta_{x_{k-1}} - 2\theta_{x_k})(x)$ so we obtain

$$\phi_h(x) = \frac{1}{h}\sum_{k=0}^{\infty} \phi(x_k)((\theta_{x_{k+1}} + \theta_{x_{k-1}} - 2\theta_{x_k})(x).$$

Applying the operator $\Pi_{t,T}$ we obtain

$$\Pi_{t,T}(\phi_h)(x) = \frac{1}{h}\sum_{k=0}^{\infty} \phi(x_k)(C_{t,T}(x,x_{k+1}) + C_{t,T}(x,x_{k-1}) - 2C_{t,T}(x,x_k))$$

so that

$$\begin{aligned} & \left| \Pi_{t,T}(\phi)(x) - \frac{1}{h}\sum_{k=0}^{\infty} \phi(x_k)(C_{t,T}(x,x_{k+1}) + C_{t,T}(x,x_{k-1}) - 2C_{t,T}(x,x_k)) \right| \\ = \quad & |\Pi_{t,T}(\phi)(x) - \Pi_{t,T}(\phi_h)(x)| \leq h e^{r(T-t)} \left\| \phi' \right\|_{\infty}. \end{aligned}$$

Let us now prove the representation formula for $\mu_{t,T}$. We suppose for a moment that $y \to C_{t,T}(x,y)$ is three times differentiable and has bounded derivatives of third order. Suppose also that ϕ is positive, differentiable and has compact support. We write

$$\begin{aligned} & \frac{1}{h}(C_{t,T}(x,x_{k+1}) + C_{t,T}(x,x_{k-1}) - 2C_{t,T}(x,x_k)) \\ & \frac{1}{h}(C_{t,T}(x,x_{k+1}) - C_{t,T}(x,x_k) - \partial_y C_{t,T}(x,x_k)h - \frac{1}{2}\partial_y^2 C_{t,T}(x,x_k)h^2) \\ & -\frac{1}{h}(C_{t,T}(x,x_k) - C_{t,T}(x,x_{k-1}) - \partial_y C_{t,T}(x,x_{k-1})h - \frac{1}{2}\partial_y^2 C_{t,T}(x,x_{k-1})h^2) \\ & +(\partial_y C_{t,T}(x,x_k) - \partial_y C_{t,T}(x,x_{k-1})) + \frac{h}{2}(\partial_y^2 C_{t,T}(x,x_k) - \partial_y^2 C_{t,T}(x,x_{k-1})). \end{aligned}$$

Since

$$\left|\frac{1}{h}(C_{t,T}(x,x_{k+1}) - C_{t,T}(x,x_k) - \partial_y C_{t,T}(x,x_k)h - \frac{1}{2}\partial_y^2 C_{t,T}(x,x_k)h^2)\right| \leq Ch^2$$

and ϕ is integrable

$$\sum_{i=0}^{\infty}\phi(x_i)\frac{1}{h}(C_{t,T}(x,x_{i+1}) - C_{t,T}(x,x_i) - \partial_y C_{t,T}(x,x_i)h - \frac{1}{2}\partial_y^2 C_{t,T}(x,x_i)h^2) \to 0.$$

For a similar reason

$$\sum_{i=0}^{\infty}\phi(x_i)\frac{h}{2}(\partial_y^2 C_{t,T}(x,x_i) - \partial_y^2 C_{t,T}(x,x_{i-1})) \to 0.$$

So we have

$$\begin{aligned}
& \lim_h \sum_{k=0}^{\infty}\phi(x_k)n(C_{t,T}(x,x_{k+1}) + C_{t,T}(x,x_{k-1}) - 2C_{t,T}(x,x_k)) \\
= & \lim_h \sum_{k=0}^{\infty}\phi(x_k)(\partial_y C_{t,T}(x,x_k) - \partial_y C_{t,T}(x,x_{k-1})) \\
= & \lim_h \sum_{k=0}^{\infty}\phi(x_k)\int_{x_{k-1}}^{x_k}\partial_y^2 C_{t,T}(x,y)dy = \int_0^{\infty}\phi(y)\partial_y^2 C_{t,T}(x,y)dy
\end{aligned}$$

and finally passing to the limit in i) we obtain

$$\Pi_{t,T}(\phi)(x) = \int_0^{\infty}\phi(y)\frac{\partial^2 C_{t,T}(x,y)}{\partial y^2}dy = \int_0^{\infty}\phi''(y)C_{t,T}(x,y)dy.$$

So we have our result for a smooth call option price. In order to obtain it for a general continuous $C_{t,T}(x,y)$ one has to employ a standard regularization procedure: for some $\varepsilon > 0$ one denotes by $\mu_{t,T}^{\varepsilon}(x,dy)$ the convolution of $\mu_{t,T}(x,dy)$ with a smooth function and then the corresponding call option prices $C_{t,T}^{\varepsilon}(x,y)$ will be smooth and so we have the above equality. Then one pas to the lit with $\varepsilon \to 0$.□

3.2 The calibration problem

We assume now that we have the data $C_{0,t_k}(x_0,K_i)$ where $0 < t_1 < ... < t_n \leq T$ and $0 < K_1 < ...K_m$ and we want to "calibrate". Since there is no underling model for the stock the volatility has no sense in this frame and so we put the problem in a more general setting: find the measures $\mu_{t,T}(x,dy), 0 \leq t \leq T, x \in R_+$ which explain the best possible the above call option prices. Let us see which are the constraints on these measures. First of all

$$(P) \quad \int_0^{\infty} 1\mu_{t,T}(x,dy) = e^{-r(T-t)} \quad \textit{and} \quad \mu_{t,T}(x,dy) \geq 0.$$

Moreover the martingale property reads

$$(M) \quad \int_0^\infty y\mu_{t,T}(x,dy) = x.$$

Finally the semi-group property reads

$$(S) \quad \int_0^\infty \int_0^\infty \phi(z)\mu_{s,t}(y,dz)\mu_{t,T}(x,dy) = \int_0^\infty \phi(y)\mu_{s,T}(x,dy), \quad \forall s < t < T, \phi \in C_+.$$

These are the basic properties.
Now, if we want these measures to feet to the data we will ask

$$(E^i_{k+1}) \quad C_{0,t_{k+1}}(x_0,K_i) = \int_0^\infty (y-K_i)\mu_{0,t_{k+1}}(x,dy) = \int_0^\infty \int_0^\infty (z-K_i)\mu_{t_k,t_{k+1}}(y,dz)\mu_{0,t_k}(x,dy).$$

The algorithm that we have in mind is evolutive with respect to the time. We assume that at step k, $\mu_{0,t_k}(x,dy)$ is known and we look for $\mu_{t_k,t_{k+1}}(y,dz)$ which satisfies $(E^i_{k+1}), i = 1,...,m$. Once we find $\mu_{t_k,t_{k+1}}$ we may produce $\mu_{0,t_{k+1}}$ using the Chapman Kolmogorov equation and we may pass to the following step of the algorithm. Except for these equations $\mu_{t_k,t_{k+1}}$ also verify the relations

$$(P_{k+1}) \quad \int_0^\infty 1\mu_{t_k,t_{k+1}}(x,dy) = e^{-r/n} \quad and \quad \mu_{t_k,t_{k+1}}(x,dy) \geq 0$$

$$(M_{k+1}) \quad \int_0^\infty y\mu_{t_k,t_{k+1}}(x,dy) = x.$$

This is an infinite dimensional problem so a discretization procedure is necessary in order to solve numerically this problem. One may imagine a parametric approach - assuming a model for $\mu_{t_k,t_{k+1}}$ - or a take a non parametric point of view as we do here. This is the subject of the following sections.

4 A finite element type algorithm

We go now further and present out calibration problem. We work in the one dimensional case, that is $D = [0,\infty)$. We assume that we are given some call option prices $\overline{C}_{0,t_k}(x_0,K_i)$ (the upper bar signals that this is the value in the experimental price table). Typically we have four epochs $t_k, k = 1,...,4$ and ten strikes $K_i, i = 1,...,10$. Using a standard linear interpolation procedure (which works in practice without any difficulty) we may extend this

table to twelve time epochs $t_k, k = 1, ..., n = 12$. So now on we assume that such a table is given.

We want to replace the semi-group of measures $\mu_{t,T}(x, dy), 0 \leq t \leq T, x, y \in (0, \infty)$ by a discrete semi-group $\pi^{ij}_{k,k+1}, k = 0, ..., n, i, j = 1, ..., M$ so that

$$\mu_{t_k,t_{k+1}}(y_i, dy) \sim \sum_{j=1}^{M} \pi^{ij}_{k,k+1} \delta_{y_j}(y),$$

where $y_i, i = 1, ..., M$ is a space grid. The strikes $K_i, i = 1, ..., 10$ will be included in the space grid but generally we can not restrict ourselves to these points. In order to obtain a sufficiently accurate approximation we need to perform our computations on a much larger grid. Typically we work with $M = 150$. Note that at time t_0 we do not have a whole grid but only one point, because the price at time zero is a deterministic known constant. So $\pi_{0,1}$ is not a matrix but just a vector $\pi^j_0, j = 1, ..., M$ so that $\mu_{0,t_1}(x_0, dy) \sim \sum_{j=1}^{M} \pi^j_{0,1} \delta_{y_j}(y)$. Note also that we may associate to these weights a Markov chain X_k so that $\pi^{ij}_{k,k+1} = P(X_{k+1} = y_j \mid X_k = y_i)$. This permits to employ the probabilistic language which is proper to this frame.

Having in mind that the stock price is expected to have an exponential behavior we choose

$$y_j = x_0 \exp(j - \frac{M}{2})h, \quad j = 1, ..., M$$

where $h > 0$ has to be chosen in such a way that, M being given, the space grid covers a significant region. This is not difficult: one has an a priory idea about the order of magnitude of the expected volatility (and so on the behavior of the queues of X_k) and then employs some elementary queues evaluations in order to obtain $P(X_k \leq x_0 \exp(-\frac{Mh}{2})) \leq \varepsilon$ and $P(X_k \geq x_0 \exp(\frac{Mh}{2})) \leq \varepsilon$ for a sufficiently small $\varepsilon > 0$. This is crucial for the practical implementation of the algorithm because this permits to handle the boundary problems.

Now our problem is to find the weights $\pi^{ij}_{k,k+1}, k = 1, ..., n = 12, i, j = 1, ..., M = 150$ which feet the best the call option prices. Following the idea from the previous section these weights have to verify the following constraints. First of all they have to give probability measures, so for every $i = 1, ..., M$

$$(P^i_k) \quad \sum_{j=1}^{M} \pi^{ij}_{k,k+1} = 1, \quad \pi^{ij}_k \geq 0.$$

They verify the martingale property that is, for every $i = 1, ..., M$

$$(M^i_k) \quad \sum_{j=1}^{M} y_j \pi^{ij}_{k,k+1} = y_i.$$

Except for the volatility we also want to compute dividends - this means that the stock gives some dividends $d_k X_k$ at the epochs t_k and these dividends are not known. We want

to compute them from the empirical price table. Note that in the presentation given in the previous sections we have implicitly assumed that there are no dividends (the dividends are null) - this was included in the martingale equation (M). If we want to treat dividends we have to replace this equation by another one which takes them into account. We have not done it before in order to simplify the presentation, but we do this now. So instead of the above equation we consider

$$(M_k^i) \quad \sum_{j=1}^{M} y_j \pi_{k,k+1}^{ij} = y_i + d_k(t_k - t_{k-1})$$

where d_k is the dividend given by the stock, at time t_k, for the period (t_{k-1}, t_k). These dividends are unknown and so except for $\pi_{k,k+1}^{ij}$ we have one more unknown at each epoch t_k.

We go further and ask to our semi-group to feet the empirical data. Having in mind the semi-group property we define by recurrence

$$\pi_{0,1}^j = \pi_{0,1}^j, \quad \pi_{0,k+1}^j = \sum_{p=1}^{M} \pi_{0,k}^p \pi_{k,k+1}^{pj}.$$

So $\pi_{0,k}^p$ represents the probability that the underlying chain starts from x_0 and arrives in y_p at time k. We also denote

$$C_{0,t_k}(x_0, y_j) = E((X_k - y_j)_+ \mid X_0 = x_0) = \sum_{p=1}^{M} (y_p - y_j)_+ \pi_{0,k}^p$$

These are the call option prices produces by the weights $\pi_{k,k+1}^{ij}$. Since we know $\overline{C}_{0,t_k}(x_0, y_j)$ for $j = j_l$ for which $y_j = K_l$, we obtain the equations

$$(E_k^{j_l}) \quad \overline{C}_{0,t_k}(x_0, K_l) = C_{0,t_k}(x_0, y_{j_l}) = \sum_{p=1}^{M} (y_p - y_{j_l})_+ \pi_{0,k}^p,$$

with $l = 1, ..., 10$. Using the Chapman Kolmogorov equation we may still write the above equations, at time t_{k+1}, as

$$(E_{k+1}^{j_l}) \quad \overline{C}_{0,t_{k+1}}(x_0, K_l) = \sum_{p=1}^{M} (y_p - y_{j_l})_+ \pi_{0,k+1}^p = \sum_{p=1}^{M} (y_p - y_{j_l})_+ \sum_{i=1}^{M} \pi_{0,k}^i \pi_{k,k+1}^{ip}.$$

Suppose that we are at the step k of our algorithm and we know from the previous step $\pi_{0,k}$ and want to compute $\pi_{k,k+1}$ - if this is done then we define $\pi_{0,k+1} := \pi_{0,k} \times \pi_{k,k+1}$ and then go to the step $k+1$. At this stage we have $2M + 10$ equations $(P_k^i), (M_k^i), i = 1, ..., M, (E_{k+1}^{i_l}), l = 1, ..., 10$ and $M \times M$ unknowns $\pi_k^{ij}, i, j = 1, ..., M$. So the problem is still

sub-determined and our basic problem now is to decide on a way or another in order handle this difficulty.

One natural idea would be to use a three branches tree (as it is done in...). This means that we suppose that for every given $i, \pi^{i,j}_{k,k+1} \neq 0$ for $j = i-1, i, i+1$ and is null otherwise. This amounts to consider a Cox Ross Ingersol type tree but with three branches instead of two - and this gives an incomplete market model and so an infinite number of risk neutral probabilities. Then we such a probability from this family which feats the best the empirical data. The system of equations is still sub-determined - we have $3M$ unknowns and only $2M+10$ equations, but clearly this problem is now much less dramatic. But from a numerical point of view this approach gives rise to instable algorithms. The reason for this is that it is extremely sensible to the geometry of the grid. Recall that we have settled an exponential space grid of exponential step $h > 0$. Then using three branches amounts to replace the log-normal law starting from a point y_i by a discrete probability concentrated in three points y_{i-1}, y_i, y_{i+1}. If the location of these points is compatible with the behavior of the sock then everything works well. But a good choice of this location suppose that we have a very good guess of the volatility and our resolution is extremely sensible to this guess. This is the reason for which we take a different way and use a finite element type approach. This approach allows to employ instead of the point itself plus two neighbors (y_i and his neighbors y_{i-1}, y_{i+1}), a much larger number of points and in some sense this has a smoothing effect.

We construct now the above mentioned trials (elements). a) We consider a standard normal random variable Δ and we take fife points $b_0 < b_1 < b_2 < b_3 < b_4$ such that $P(\Delta < b_0) = P(\Delta > b_4) = 1/100$ and $P(b_p < \Delta < b_{p+1}) = \frac{1}{4} \times (1 - \frac{2}{100}), p = 0, 1, 2, 3$. Of course we will have $b_2 = 0, b_1 = -b_3, b_0 = -b_4$. b) We compute the implicit volatilities $\overline{\sigma}^{j_l}_k$ corresponding to the experimental call option prices $\overline{C}_{0,t_{k+1}}(x_0, K_l), l = 1, ..., 10$ and then we use a linear interpolation in order to produce all the $\overline{\sigma}^i_k, i = 1, ..., M$. c) Then we define $a_p = \overline{\sigma}^i_k \times \sqrt{1/12} b_p$ (recall that the time step is $\delta = 1/12$). The significance of $a_p, p = 0, ..., 4$ is simple: they represents the points which divide in equal parts the total mass of the probability density of a normal distributed random variable of variance $\overline{\sigma}^i_k \times \sqrt{1/12}$. This choice appears as natural if one supposes that $\mu_{t_k, t_{k+1}}(x, dy)$ is the law of a random variable of the form $\exp(\overline{\sigma}^i_k \times \sqrt{1/12})\Delta + \frac{1}{2}(\sigma^i_k)^2/12)$ - as it would be the case in the Black Scholes model.

We take ϕ_p to be the piecewise linear function such that $\phi_p(a_p) = 1, \phi_p(z) = 0$ if $z \in [a_{p-1}, a_{p+1}]^c, p = 1, 2, 3$. We will achieve a finite element method for each starting point y_i, so we have to center our trials in this point. So our trials will be $\phi_p(\ln \frac{y}{y_i})$.

Now we think to $\pi^{ij}_{k,k+1}$ to be a function in the forward argument, that is $\pi^{ij}_{k,k+1} = \pi^i_{k,k+1}(y_j)$ and project it on the trials

$$(*) \quad \pi^{ij}_{k,k+1} = \pi^i_{k,k+1}(y_j) = \sum_{p=1}^{3} \lambda^i_p \phi_p(\ln \frac{y_j}{y_i}).$$

Remark 4.1 *As we mentioned in the beginning we make no explicit model hypothesis. But at this stage it appears clearly that an a priory guess about the underlying dynamics is implicit*

in our algorithm. In fact the choice of the space grid and of the trials presented above suppose a certain idea about the geometry of the problem. The fact that we expect that the underlying stock has an exponential type dynamics appears in the choice of an exponential space grid. Choosing the trials as we have done suppose also that the underlying stock follows continuous trajectories. If we would want to reproduce a jump type dynamics we would have to use other trials - maybe larger or maybe located at a certain distance from the starting point y_i. So an a priory model guess comes on here. But this is a very flexible way of including model hypothesis in the algorithm and a large class of variants may be treated in this frame. The only rigid hypothesis concerns Markovianity which is implicit in the fact that the price of an European option at time t depends on the price of the underlying stock at this moment only.

The question about the "model hypothesis" may be asked in another way also. It is known that as long as we have informations on a finite number of epochs only, one is not able to "precise the model": a Dupire model or a jump model may explain the call option prices as well (see Rama Cont and ...[??] for a detailed discussion and numerical experiments about this matter). Then the question is: the finite elements method presented here will choose the Dupir model or the jump model? The answer is that the implicit choice depends on the geometry of the trials that one employs. Trials concentrated around the starting point will give results which are close to Dupire's model and large trials (or trials located far from the starting point) will give a results closed to the jump model. And there is a flexibility because we are not obliged to decide that we work with Dupire's model, pure jumps models or a mixing of the two ones.

We are now ready to present our algorithm. The initialization step $k = 0$ is different from the current step and we postpone it to the end. We suppose now that the step $k-1$ is already achieved and then we have the weights $\pi^j_{0,k} = P(X_k = y_j \mid X_0 = x_0), j = 1, ..., M$.

STEP k. The equations are the following. For each $i = 1, ..., M$ we have the equations

$$(P^i_k) \quad 1 = \sum_{j=1}^{M} \pi^{ij}_{k,k+1} = \sum_{j=1}^{M}\sum_{p=1}^{3} \lambda^i_p \phi_p(\ln \frac{y_j}{y_i}) = \sum_{p=1}^{3} \lambda^i_p \left(\sum_{j=1}^{M} \phi_p(\ln \frac{y_j}{y_i}) \right),$$

$$(M^i_k) \quad y_i + d_k\delta = \sum_{j=1}^{M} y_j \pi^{ij}_{k,k+1} = \sum_{j=1}^{M} y_j \sum_{p=1}^{3} \lambda^j_p \phi_p(\ln \frac{y_j}{y_i})$$
$$= \sum_{p=1}^{3} \lambda^j_p \left(\sum_{j=1}^{M} y_j \phi_p(\ln \frac{y_j}{y_i}) \right).$$

We solve these first two equations explicitly.. We denote

$$\alpha^i_p = \sum_{j=1}^{M} \phi_p(\ln \frac{y_j}{y_i}), \quad \beta^i_p = \sum_{j=1}^{M} y_j \phi_p(\ln \frac{y_j}{y_i}).$$

Form the first equation we obtain

$$\lambda_3^i = \frac{1 - \lambda_1^i \alpha_1^i - \lambda_2^i \alpha_2^i}{\alpha_3^i}$$

and from the second equation we obtain

$$y_i + d_k\delta = \lambda_1^i \beta_1^i + \lambda_2^i \beta_2^i + \frac{1 - \lambda_1^i \alpha_1^i - \lambda_2^i \alpha_2^i}{\alpha_3^i}\beta_3^i$$

which gives

$$y_i - \frac{\beta_3^i}{\alpha_3^i} = \lambda_1^i(\beta_1^i - \frac{\alpha_1^i}{\alpha_3^i}\beta_3^i) + \lambda_2^i(\beta_2^i - \frac{\alpha_2^i}{\alpha_3^i}\beta_3^i)$$

and finally

$$\begin{aligned}
\lambda_2^i &= \frac{(y_i + d_k\delta - \frac{\beta_3^i}{\alpha_3^i}) - \lambda_1^i(\beta_1^i - \frac{\alpha_1^i}{\alpha_3^i}\beta_3^i)}{\beta_2^i - \frac{\alpha_2^i}{\alpha_3^i}\beta_3^i} \\
&= \frac{((y_i + d_k\delta)\alpha_3^i - \beta_3^i) - \lambda_1^i(\beta_1^i\alpha_3^i - \alpha_1^i\beta_3^i)}{\beta_2^i\alpha_3^i - \alpha_2^i\beta_3^i} \\
&= \gamma_2^i(d_k) - \lambda_1^i\mu_2^i
\end{aligned}$$

with

$$\gamma_2^i(d_k) = \frac{(y_i + d_k\delta)\alpha_3^i - \beta_3^i}{\beta_2^i\alpha_3^i - \alpha_2^i\beta_3^i}, \quad \mu_2^i = \frac{\beta_1^i\alpha_3^i - \alpha_1^i\beta_3^i}{\beta_2^i\alpha_3^i - \alpha_2^i\beta_3^i}.$$

By symmetry we obtain a similar expression for λ_3^i

$$\lambda_3^i = \gamma_3^i(d_k) - \lambda_1^i\mu_3^i$$

with

$$\gamma_3^i(d_k) = \frac{(y_i + d_k\delta)\alpha_2^i - \beta_2^i}{\beta_3^i\alpha_2^i - \alpha_3^i\beta_2^i}, \quad \mu_3^i = \frac{\beta_1^i\alpha_2^i - \alpha_1^i\beta_2^i}{\beta_3^i\alpha_2^i - \alpha_3^i\beta_2^i}.$$

We come now to the equations (E_k^i). We employ the semi-group equation and write

$$\begin{aligned}
(E_k^j) \quad C_{0,t_{k+1}}(x_0,y_j) &= \sum_{l\geq j}(y_l-y_j)\pi^j_{0,k+1} = \sum_{l\geq j}(y_l-y_j)\sum_{i=1}^M \pi^{il}_{k,k+1}\pi^i_{0,k} \\
&= \sum_{l\geq j}(y_l-y_j)\sum_{i=1}^M \pi^i_{0,k}\sum_{p=1}^3 \lambda^i_p\phi_p(\ln\frac{y_l}{y_i}) \\
&= \sum_{i=1}^M\sum_{p=1}^3 \lambda^i_p \sum_{l\geq j}(y_l-y_j)\pi^i_{0,k}\phi_p(\ln\frac{y_l}{y_i}) \\
&= \sum_{i=1}^M(\lambda^i_1+\gamma^i_2(d_k)-\lambda^i_1\mu^i_2+\gamma^i_3(d_k)-\lambda^i_1\mu^i_3) \\
&\quad \times\sum_{l\geq j}(y_l-y_j)\pi^i_{0,k}\phi_p(\ln\frac{y_l}{y_i}) \\
&= \sum_{i=1}^M \lambda^i_1(1-\mu^i_2-\mu^i_3)\sum_{l\geq j}(y_l-y_j)\pi^i_{0,k}\phi_p(\ln\frac{y_l}{y_i}) \\
&\quad +\sum_{i=1}^M(\gamma^i_2+\gamma^i_3)(d_k)\sum_{l\geq j}(y_l-y_j)\pi^i_{0,k}\phi_p(\ln\frac{y_l}{y_i}).
\end{aligned}$$

Since we know the values of $C_{0,t_{k+1}}(x_0,y_j)$ for $y_j = K_l, l = 1,...,10$ we have here a system of 10 equations with $M = 150+1$ unknowns $\lambda^i_1, i = 1,...,M$ and d_k. This system of linear equations is still un-determined. Our first attempt was to contouring this difficulty using interpolation. By a more or less sophisticated method one interpolates and produces $\overline{C}_{0,t_{k+1}}(x_0,y_j)$ for every $j = 1,...,M$ and not only for $y_j = K_l, l = 1,...,10$. But this does not work. A first idea would be that this is because the interpolation introduces errors - but in fact we checked that the interpolation was very accurate, and moreover, we performed numerical experiments in which we gave directly all the $\overline{C}_{0,t_{k+1}}(x_0,y_j), j = 1,...,M$ produced by our theoretical model - and this does not work also. So the interpolation error is not the reason for which this approach fails. The real reason (numerical evidence) is that even if we have a system of 150 equations with 150 unknowns which is theoretically well determined, we may produce two different volatilities σ and σ' which are significantly different but such that the corresponding call prices $\overline{C}_{0,t_{k+1}}(x_0,y_j)$ and $\overline{C}'_{0,t_{k+1}}(x_0,y_j)$ are extremely close each other. So it turns out that the call option prices are not sufficiently sensible in order to distinguish between volatilities - at list from a numerical point of view. So we have to change the "scale" in which we work by a more significant one, and the natural idea is to use implied volatilities.. So we live out the resolution of the above linear system of equations and focus on implied volatilities.. This leads us to solve the following non linear optimization problem.

For each $\lambda = (\lambda^1, ..., \lambda^M) \in R^M$ and $d_k \geq 0$ we compute

$$C^j(\lambda, d_k) \quad := \sum_{i=1}^{M} \lambda^i (1 - \mu_2^i - \mu_3^i) \sum_{l \geq j} (y_l - y_j) \pi_{0,k}^i \phi_p(\ln \frac{y_l}{y_i})$$
$$+ \sum_{i=1}^{M} (\gamma_2^i + \gamma_3^i)(d_k) \sum_{l \geq j} (y_l - y_j) \pi_{0,k}^i \phi_p(\ln \frac{y_l}{y_i}).$$

Then we think that $C^j(\lambda, d_k) = C_{0,t_{k+1}}(x_0, y_j)$ so it represents the value of a call option (λ^i play the part of λ_1^i). We denote by $Iv_j(\lambda, d_k)$ the implied volatility associated to this call option price. We also assume that $y_j = K_l$ (that is $j = i_l$) and we compute the implied volatility $\overline{Iv}_l$ corresponding to the experimental call option price $\overline{C}_{0,t_{k+1}}(x_0, K_l)$. Note that here we have a problem because when computing the implied volatilities we have to take care of the dividends.. This means that the log normal distribution function that we inverse contains d_k - and the interest rate r as well. But this is not a difficult problem: one just has to multiply first with $\exp((-r + \sum_{i=1}^{k} d_i)t_k)$. Note that $d_i, i = 1, ..., k-1$ are already known but d_k is unknown and it appears in the multiplication both for $C^j(\lambda, d_k)$ and $\overline{C}_{0,t_{k+1}}(x_0, K_l)$. So $\overline{Iv}_l = \overline{Iv}_l(d_k)$.

The first quantity that we want to minimize is

$$A(\lambda) := \sum_{l=1}^{10} \left| Iv_{j_l}(\lambda, d_k) - \overline{Iv}_l(d_k) \right|^2.$$

So we do not ask that the prices produced by our semi-group are closed to the market prices (which gives an un-determined linear system) but that the implied volatilities are closed. It turns out that this is the correct scale in which the problem has to be settled - if we use not this scale but directly the price table our algorithm works much worse (experimental evidence).

Remark 4.2 *Once again we would ask if our algorithm is model free or not - because we are using the implicit volatility which is proper to the Black Scholes model. But notice that the implicit volatility is used just as a "distance" which measures the fact that we are more or less close to the empirical data.*

The second quantity that we want to minimize is the distance between $\lambda = \lambda_1$ and λ_2, λ_3 where

$$\lambda_2^i = \gamma_2^i(d_k) - \lambda^i \mu_2^i, \quad \lambda_3^i = \gamma_3^i(d_k) - \lambda^i \mu_3^i.$$

This avoids to have all the mass in λ_1^i. The corresponding coast function is

$$B(\lambda) = \sum_{i=1}^{M} (\left| \lambda^i - \lambda_2^i \right|^2 + \left| \lambda^i - \lambda_3^i \right|^2).$$

Now the coast function which we minimize is

$$C(\lambda) = A(\lambda) + pB(\lambda)$$

where p is a parameter. In our concrete algorithm we take $p = 10/M$.

We look for the $\lambda^* = argmin\ C(\lambda)$, under the constraint $\pi^{ij}_{k,k+1} \geq 0, i, j = 1, ..., M$. We use a quasi-Newton algorithm which is already implemented in Scilab. The starting point of the optimization algorithm is the value of λ which has been founded at the previous step of the algorithm.. This ensures a certain stability in time. We are able to include in our coast function a constraint concerning stability in space - the fact that $i \to \lambda_i$ does not move very fast. But we at this stage we use no such a constrained and the algorithm remains stable anyway.

Once $\lambda = (\lambda^1, ..., \lambda^M)$ is computed we put $\lambda^i_1 = \lambda^i$, compute λ^i_2, λ^i_3 and use $(*)$ in order to obtain $\pi^{ij}_{k,k+1}$. Then we use the Chapman Kolmogorov equation in order to compute $\pi_{0,k+1}$ and we are ready for the following step.

STEP 0. We recall that at step zero we compute $\pi^j_{0,1} = P(X_1 = y_j \mid X_0 = x_0), j = 1, ..., M$. Here the degree of indeterminacy is much less important because we have only M unknowns. This is why we will employ a much more important number of finite elements, namely 64. This is also necessary because we need a very accurate result at this stage: an important error would orient the algorithm in a bad direction. The construction of the points $b_i, i = 0, ..., 64$ is similar: we first take the points which cut the mass of the standard normal density in 65 equal parts. Then we construct $a_i, i = 1, ..., 64$ by normalization with $\sigma\sqrt{\delta}$ with $\delta = 1/12$ and σ is the implied volatility at the money, given by $C_{0,t_1}(x_0, x_0)$. So we assume that the functions $\phi_p, p = 1, ..., 64$ are now given (the same construction as before) and we use the trials $\phi_p(\ln \frac{y}{x_0})$. As before we write the two equations (P_0) and (M_0) (note that we have one equation of each type) and then we write down the equations $(E^{j_l}_0), l = 1, ..., 10$ which give:

$$(**) \quad C_{0,t_1}(x_0, K_l) = C_{0,t_1}(x_0, y_{j_l}) = \sum_{p=1}^{64} \lambda_p \sum_{r \geq j_l} (y_r - y_{j_l}) \pi^r_{0,1} \phi_p(\ln \frac{y_r}{x_0}).$$

In this case we do not use (P_0) and (M_0) in order to eliminate two variables - we will keep these equations as constraints. Note that at time zero there are no dividends.. Now we consider the coast function $C(\lambda) = A(\lambda) + pB(\lambda)$ with $\lambda = (\lambda_1, ..., \lambda_{64})$

$$A(\lambda) := \sum_{l=1}^{10} \left| Iv_{j_l}(\lambda) - \overline{Iv}_l \right|^2, \quad B(\lambda) = \sum_{p=1}^{64} |\lambda_p - \lambda_{p+1}|^2.$$

Here $Iv_{j_l}(\lambda)$ is the implied volatility corresponding to $C_{0,t_1}(x_0, K_l)$ computed with our λ and $\overline{Iv}_l$ is the implied volatility associated to the experimental value $\overline{C}_{0,t_1}(x_0, K_l)$.

4.1 Computation of the volatility and of the dividends

Our algorithm does not depend on the concept of volatility - if not by the geometry of the grid and of the trials. But even for this we do not use some values of the volatility produced inside the algorithm but the implied volatilities associated to the experimental price table. And the table of experimental implied volatilities represent a sufficiently good guess for our purpose.. So we may live out the problem of computing volatilities. But this is the usual language for people working in finance and so it seems useful to produce the volatility table which is naturally associated to the semi-group that we have already computed. As it is clear from numerical experiments the table σ_k^i that we produce is significantly different from the experimental implied volatility table $\overline{\sigma}_k^i$ and, as long as we use synthetic data our volatility surface is much closer to the real volatility then the implied volatility - which means that some work has been done. We may also think that the difference between σ_k^i and the precise volatility represents a good error measure..

Except for the volatility we also want to compute dividends - this means that the stock gives some dividends $d_k X_k$ at the epochs t_k and these dividends are not known. We want to compute them from the empirical price table.

The concept of volatility has no sense in the abstract setting that we used up to now, so in order to define the volatility we have to consider a Black Scholes type dynamics for the stock. We do it at the level of the Markov chain $X_k, k = 0, ..., n = 12$ which is associated to the discrete semi-group $\pi_{k,k+1}$. We assume that under the risk neutral probability

$$X_{k+1} = X_k + \sigma_k(X_k)\Delta_k\sqrt{\delta} + d_k X_k \delta$$

where $\delta = t_{k+1} - t_k = \frac{1}{12}$ and $\Delta_k, k = 0, ..., n-1$ is a sequence of standard normal distributed random variables. As we mentioned above d_k represents the dividends given by the stock X_k at time t_k. Both $\sigma_k(x)$ and d_k are unknowns and they have to be computed from the $\pi_{k,k+1}^{ij}$ obtained before..

We write first $E(X_{k+1} - X_k \mid F_k) = E(X_{k+1} - X_k \mid X_k) = d_k X_k \delta$ which gives

$$d_k = \frac{E(X_{k+1} - X_k \mid X_k)}{X_k \delta}.$$

This formula has to work on each set $\{X_k = y_i\}$ so we have for every $i = 1, ..., M$

$$d_k = \frac{E(X_{k+1} - X_k \mid X_k = y_i)}{y_i \delta} = \frac{1}{y_i \delta} \sum_{j=1}^{M} (y_j - y_i)\pi_{k,k+1}^{ij}.$$

Note that we will obtain the same value of d_k does not meter the value of i. This is because we have put this condition in the equation (M_i) already. So if some differences appear this is due to some numerical errors coming on in our algorithm - and then we have to use some standard projection argument which gives the value of d_k which feats the best all the equations. At the contrary, if one considers that d_k is allowed to depend on the position of

X_k (which does not seem natural from an economic point of view) then one has to change the formulation of the problem, namely of the equation (M_i).

Let us compute $\sigma_k(y_i)$. We write $X_{k+1} - X_k - d_k X_k \delta = \sigma_k(X_k)\Delta_k\sqrt{\delta}$ and taking conditional expectations we obtain $E(|X_{k+1} - X_k - d_k X_k \delta|^2 \mid X_k = y_i) = y_i^2 \sigma_k^2(y_i)\delta$ which gives

$$\sigma_k^2(y_i) = \frac{E(|X_{k+1} - X_k - d_k X_k \delta|^2 \mid X_k = y_i)}{y_i^2 \delta} = \frac{1}{y_i^2 \delta} \sum_{j=1}^{M} (y_j - y_i - d_k y_i \delta)^2 \pi_{k,k+1}^{ij}.$$

So this is the volatility which is naturally associated to $\pi_{k,k+1}^{ij}$. In our algorithm we have produced a version $\widetilde{\sigma}_k(y_j)$ which represents a more stable version. This version is produced using the following optimization algorithm.. We consider the coast function

$$C(\sigma) = \sum_{i=1}^{M} \left|\sigma^i - \sigma_k(y_i)\right|^2 + q \sum_{i=1}^{M} \left|\sigma^i - \widetilde{\sigma}_{k-1}(y_i)\right|^2$$

with $\sigma = (\sigma^1, ..., \sigma^M)$ and q a real number (in our concrete computations we took $q = 1/M = 1/150$). $\widetilde{\sigma}_{k-1}$ is the smoothed volatility computed at the previous step and $\widetilde{\sigma}_0$ is just the experimental implied volatility at the money..

It turns out that $\widetilde{\sigma}_k$ is closer to the real value of the volatility if we consider synthetic data.

5 Numerical experiments

In order to test our algorithm, we try to calibrate empirical datas created synthetically from known models. We focus essentially on Dupire model with four types of volatility:

- σ constant: $\sigma = 0.3$. This is the Black Scholes model denoted *BS*.
- $\sigma(t,x) = 15/x$ denoted *Brow*.
- $\sigma(t,x) = 0.05 + 0.1 * \exp(-x/100) + 0.01 * t$ denoted *Voltx*.
- $\sigma(t,x) = 0.3 * \mathbf{1}_{x\in[90,110]} + 0.15 * \mathbf{1}_{x\notin[90,110]}$ denoted *Jump*.

All the experiments are done with a starting point set at $S_0 = 100$, an interest rate equal to 0, and a maximal maturity equal to 1. We take 150 points of discretization in space and 12 points in time. Recall that this the "numerical grid" on which we work. The grid on which the experimental data is given is much more poor (see section 4).

Remark 5.1 *A parameter is very important for our algorithm: the extreme value of the grid. To determine it, we have an heuristic approach. We first set a very small grid. If at a time step, the call are not fitted we enlarge the grid. At the first value the call are fitted, the extreme value are found.*

5.1 Black Scholes model

model	Number of datas	Time step	Volatility error	Volatility error ($K \in [70, 130]$)	Volatility error ($K \in [80, 120]$)	Call error	Put error
BS	20	2	20	2.3	2.2	10^{-7}	2.1
BS	20	6	10	2.1	1.9	10^{-7}	1.3
BS	20	12	5	1.4	1.2	10^{-7}	1
BS	10	2	20	4.8	4.8	10^{-7}	2.6
BS	10	6	10	1.4	1.6	10^{-7}	1.7
BS	10	12	10	1.1	1.3	10^{-7}	1
BS	5	2	20	4.1	3.5	10^{-7}	2.4
BS	5	6	10	4	4	10^{-7}	1.7
BS	5	12	10	1.6	1.5	10^{-7}	1
BROW	5	2	25	9.7	5.9	10^{-7}	3.4
BROW	5	6	20	5.7	1.8	10^{-7}	2.8
BROW	5	12	15	2.4	1.3	10^{-7}	1.6
Voltx	5	2	XX	50	28	10^{-7}	3.6
Voltx	5	6	20	3.9	3.3	10^{-7}	2.2
Voltx	5	12	20	1.4	1	10^{-7}	1.1

Table 1: Precision of the algorithm. The error are in percent

We use for these test a volatility equal to 0.3. We use either $5, 10$ or 20 datas generated by the closed formulas. The table 1 illustrates the evolution of the error at several steps of the algorithm, with different number of data and different underlying models. The error is measured in several ways. First, we consider the error between the theoretical volatility and the volatility produced by our algorithm. We take a mean value over all the points in the grid first and then on the strike in the center of the grid: $K \in [80, 120]$, $K \in [70, 130]$. It is clear that the results are much more better on the center and the big errors are done on the border. Finally, in the last two columns we give the mean value for the call (resp. put) prices. The mean value is taken on $K \in [80, 120]$. The error for the call option is practically equal to zero (we fit the datas) but for the put option this error becomes sgnificant. Anyway it remains at a good level.

Note that (look at the first nine lines) as the algorithm evolutes (step 2 $\rightarrow$ step 6 $\rightarrow$ step 12) the error decreases significantly. It is natural because we take into account more and more datas.

The sensibility to the number of datas (20,10 or 5) at each tile level seems not very important.

Finally we look to the last line concerning "Voltx". Note that the errors are almost the same as for the "BS" model so the algorithm seems to reproduce well any shape of local volatility.

Of course the main interest of an algorithm of calibration is that the experimental data must be fitted precisely. In the figure 1, we plot the Call obtained by our algorithm w.r.t. the theoretical value. We also compute at each time step, the values of the Put option.

This is a test significance because it shows that our algorithm is able to price other options without using the volatility. We can observe on this figure that at step 2, the call and the put are well fitted. We go a step further and look to the volatility: in figure 2 we plot four curves:

- The real volatility which is constant.
- The approached volatility obtained by calibration.
- The real implied volatility which is equal to the real volatility in the Black and Scholes model.
- The numerical implied volatility (the implied volatility computed with the numerical call prices).

Figure 2 shows that at step two, the implied volatility is perfectly fitted on the center but becomes bad on the borders. This seems natural because we have done just one time step. So, roughly speaking the underlying diffusion has not the tile to go very far from the starting point. In some sense, we have still no information far from the starting point. In contrast with this, at step 12 (figure 4) the implied volatility is perfectly fitted on all the space grid.

We look now to the real volatility and to the numerical approximation of this volatility. The result is significantly less goood as for the implied volatility but it remains at the level of 0.4% on the center. Compose with the analogous result at step 12 (figure 4); Here the numerical volatility is much more stable in space and convenient error is obtained.

Another interesting question is how closed are the real probability density and the probability density produced by our algorithm. This is explain in the next two graph who plot the real density and ours in the center (for the figure 5) and at the extreme strike (for the figure 6). We can see in these figures that the density is well approached by our algorithm even at step 2.«z

At step 6 and 12 (see figure 3,4), the volatility is fitted better on the hole grid, and the call, put and densities are still good. For the Black Scholes model we also want to see if other options are well computed. So we use our semi-group to approach the price of 9 American put of strike $80, 85, 90, 95, 100, 105, 110, 115, 120$. We use the semi-group obtained with 5,10 and 20 experimental datas. The results are good, the precision is around 1%. These results are plotted in the figure 7 and are almost the same for 20,10 or 5 experimental datas. Finally we use the semi-group produces by our algorithm in order to compute the price of a barrier option (see figure 8). The results are rather bad. But note that the algorithm that we use is rather rough and does not take into account the specification of such an option. And such an algorithm gives bad results even if we know the precise underlying semi-group.

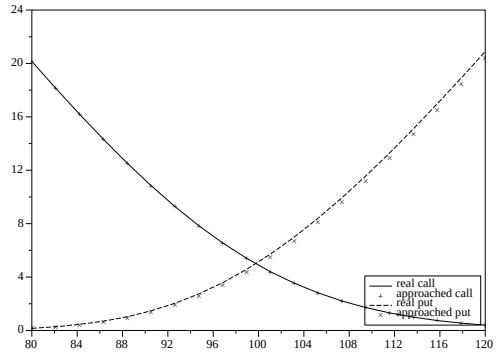

Figure 1: Computation of the Call and the Put in the Black Scholes model: time step = 2, 20 experimental datas.

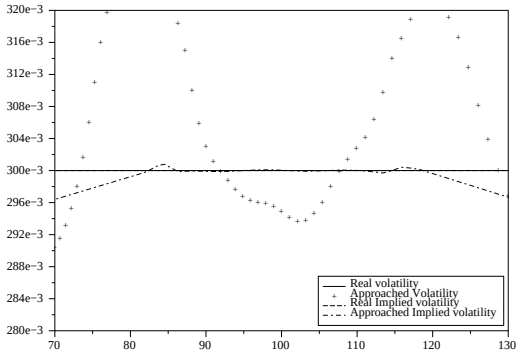

Figure 2: Computation of the volatility in the Black Scholes model: time step = 2, 20 experimental datas.

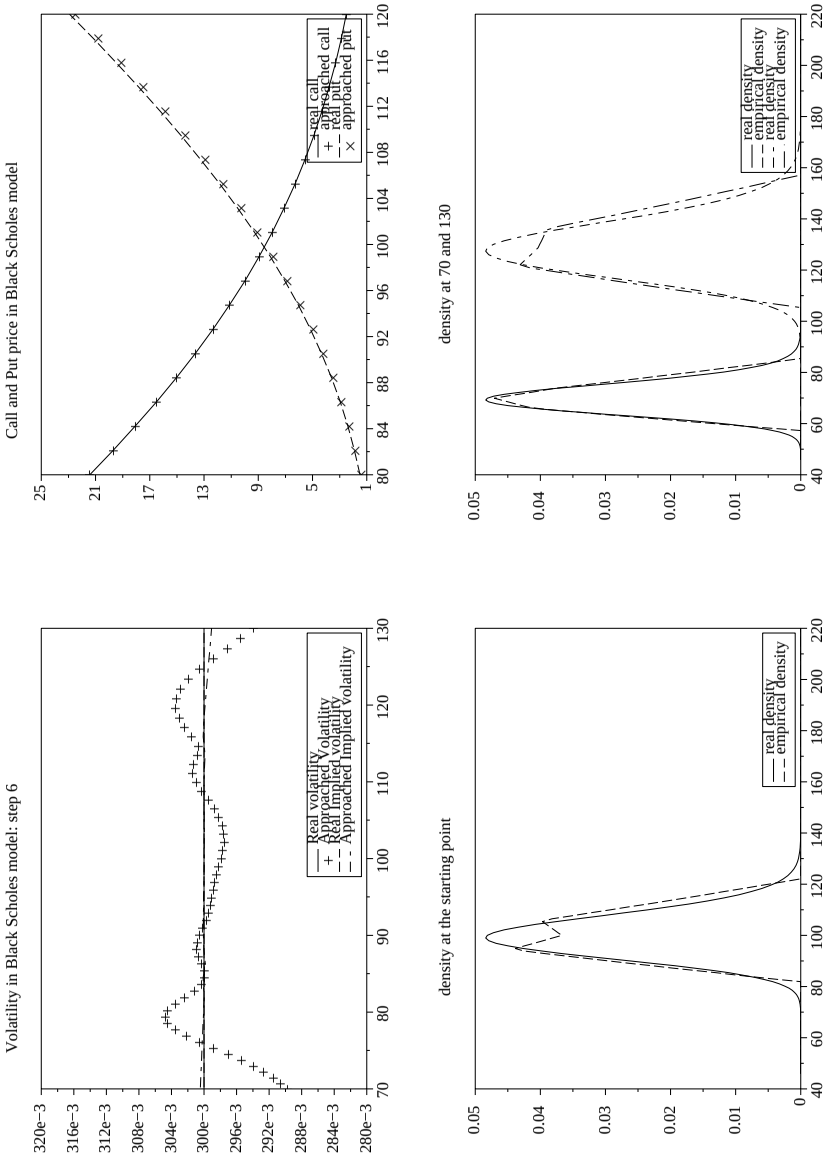

Figure 3: Computation in the Black Scholes model: time step = 6, 20 experimental datas.

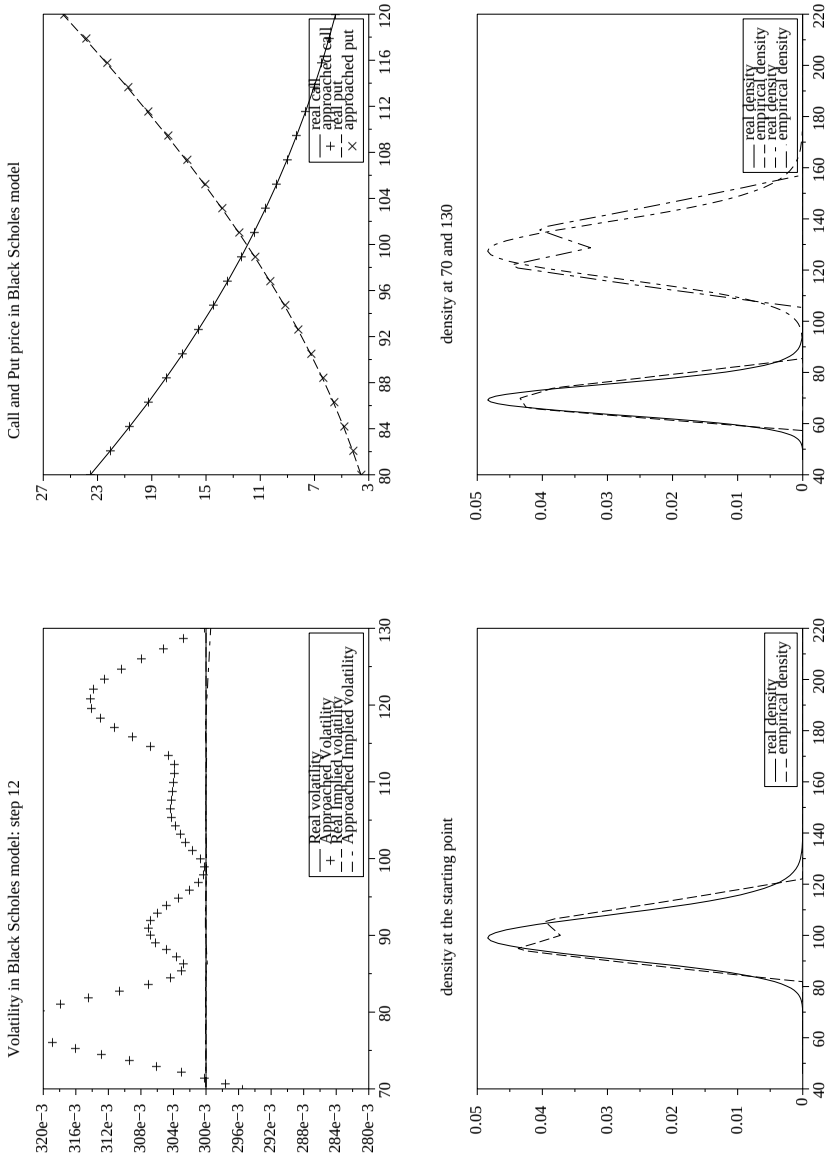

Figure 4: Computation in the Black Scholes model: time step = 12, 20 experimental datas.

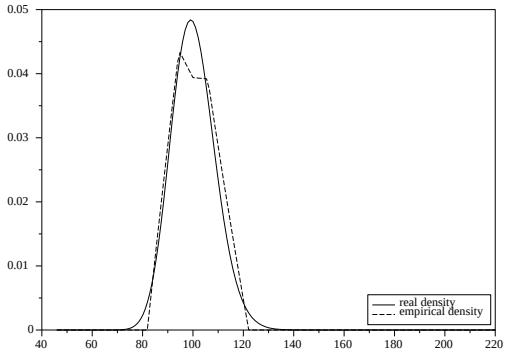

Figure 5: Computation of the density for $x = S_0$: time step $= 2$, 20 experimental datas.

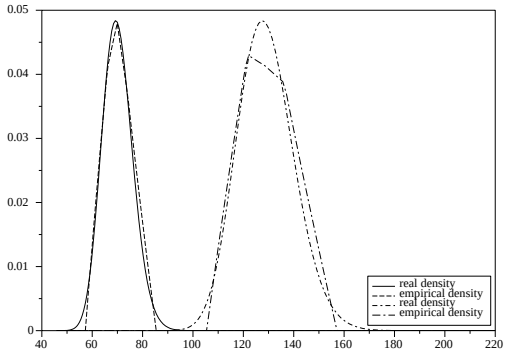

Figure 6: Computation of the density for $x = 80$ and $x = 120$: time step $= 2$, 20 experimental datas.

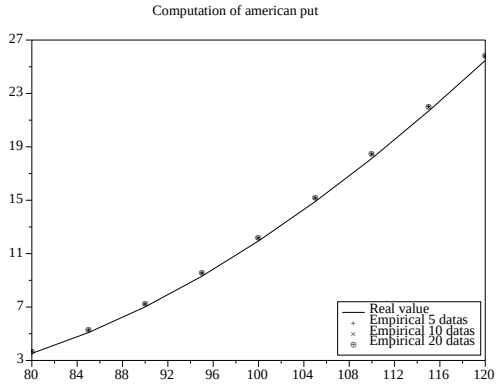

Figure 7: Computation in the Black Scholes model: time step = 12, 20 experimental datas.

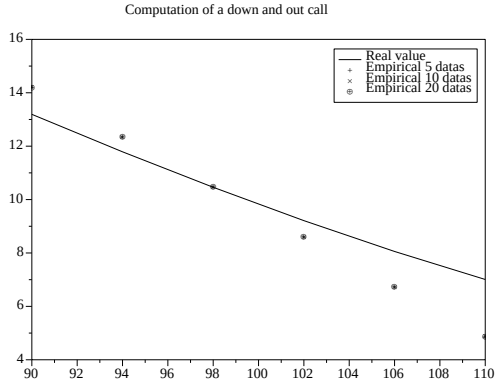

Figure 8: Computation in the Black Scholes model: time step = 12, 20 experimental datas.

5.2 Dupire model

The three other volatilities that we use enter in the frame of Dupire's model. We give the graphs corresponding to Voltx (step 12, figure 9) and Jump (step 6, figure 10). In both cases the volatility smile is perfectly fitted in the center. We have also a good shape for the real volatility in voltx. In the case of a jump of volatility, we see that the numerical approximation is sensible to the jump but gives a regularized version of the real shape. In both case, the put options are well computed.

6 Heston model

In the Heston model, the underlying asset follows the stochastic differential equation:

$$dS_t = rS_t dt + \sqrt{v_t} S_t dW_t^1$$
$$dv_t = k(\theta - v_t)dt + \sigma\sqrt{v_t} dW_t^2,$$

where W_t^1 and W_t^2 are two correlated brownian motion with $\langle W^1, W^2\rangle_t = \rho t$. This model has a stochastic volatility. Thus, it is not a markovian model which is one of our assumption. Theoretically, we cannot calibrate such model. We proceed as follows:

- We take call options prices generated by the closed formula in the Heston model ($r = 0$, $k = .01$, $\theta = 2$).
- We run our algorithm to obtained a semi-group.
- We compute put options and compare it to the real prices.

In the figure 11 and 12, the first graph presents the call and the put options in their real scale. The second one presents them in the implied volatility scale.

$\sigma = 0.2$ This value for the volatility of the volatility is low, so we are close to a Dupire model. We observe in figure 12 that the volatility smile is fitted perfectly. The error on the put options are still acceptable. We remark that the put options are fitted better for strike in the center than on the border.

$\sigma = 1$ This is an extreme value, it implies that the process v_t has high perturbation. In fact, our algorithm does not succeed in calibrate this, and the put error is over 10%. Even the call option are not well fitted. But, we have to see that the smile is very irregular and this is not real in practice.

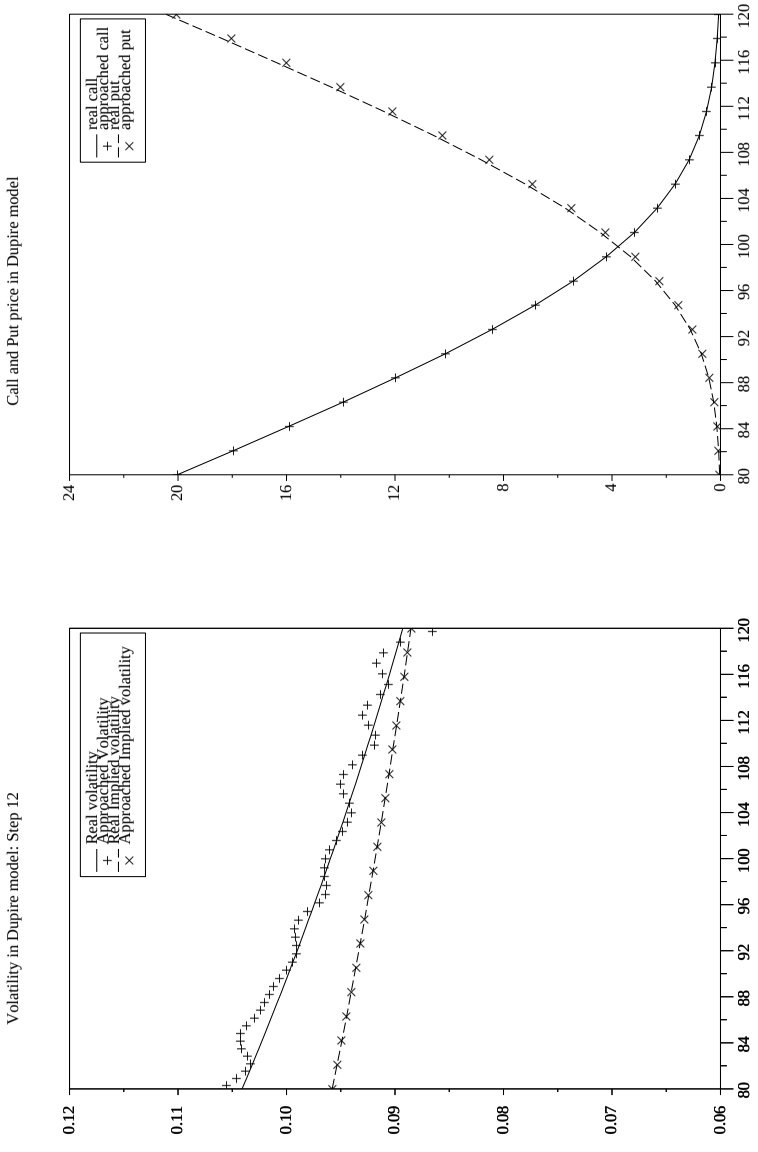

Figure 9: Computation in Voltx model: time step = 12, 10 experimental datas.

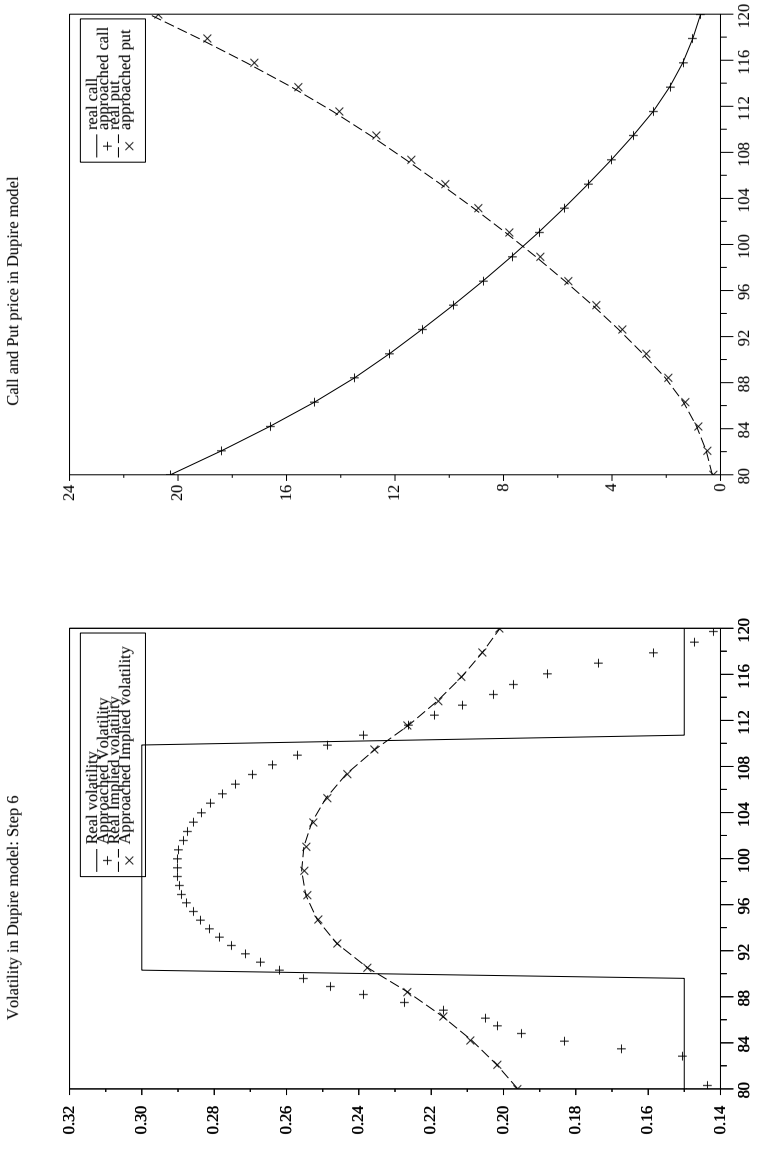

Figure 10: Computation in "Jump" model: time step = 6, 20 experimental datas.

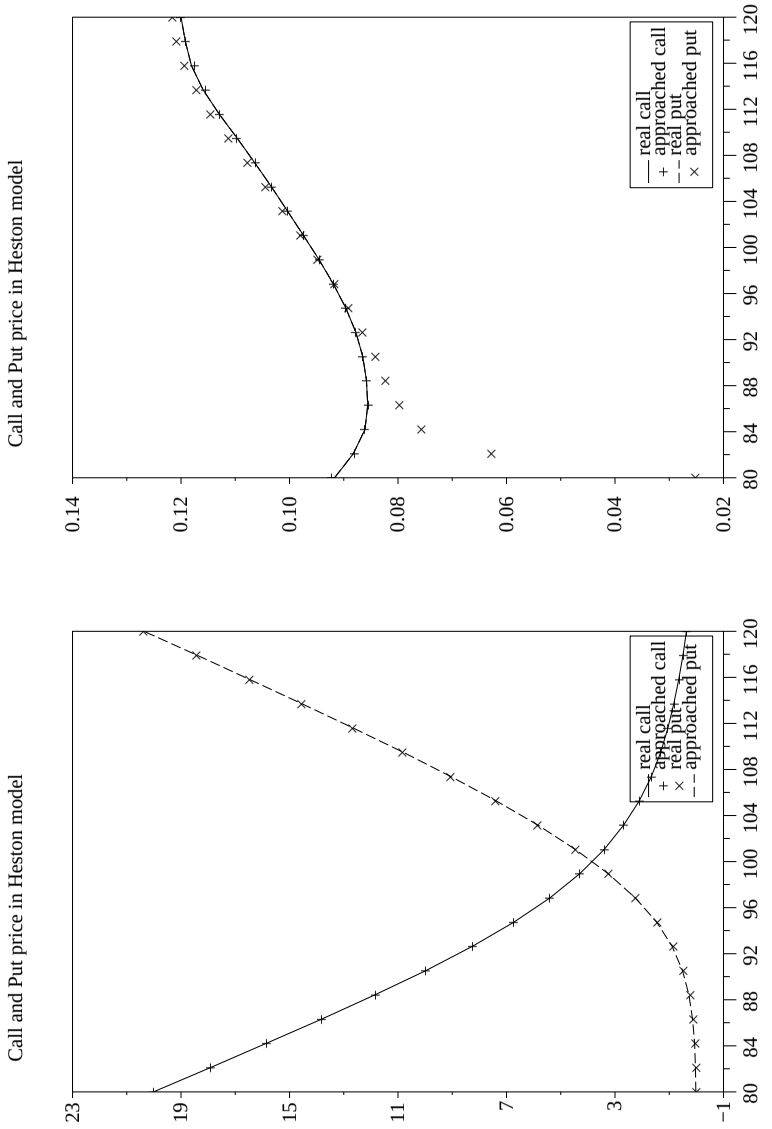

Figure 11: Computation in Heston model: time step = 12, 20 experimental datas, $\sigma = 0.2$.

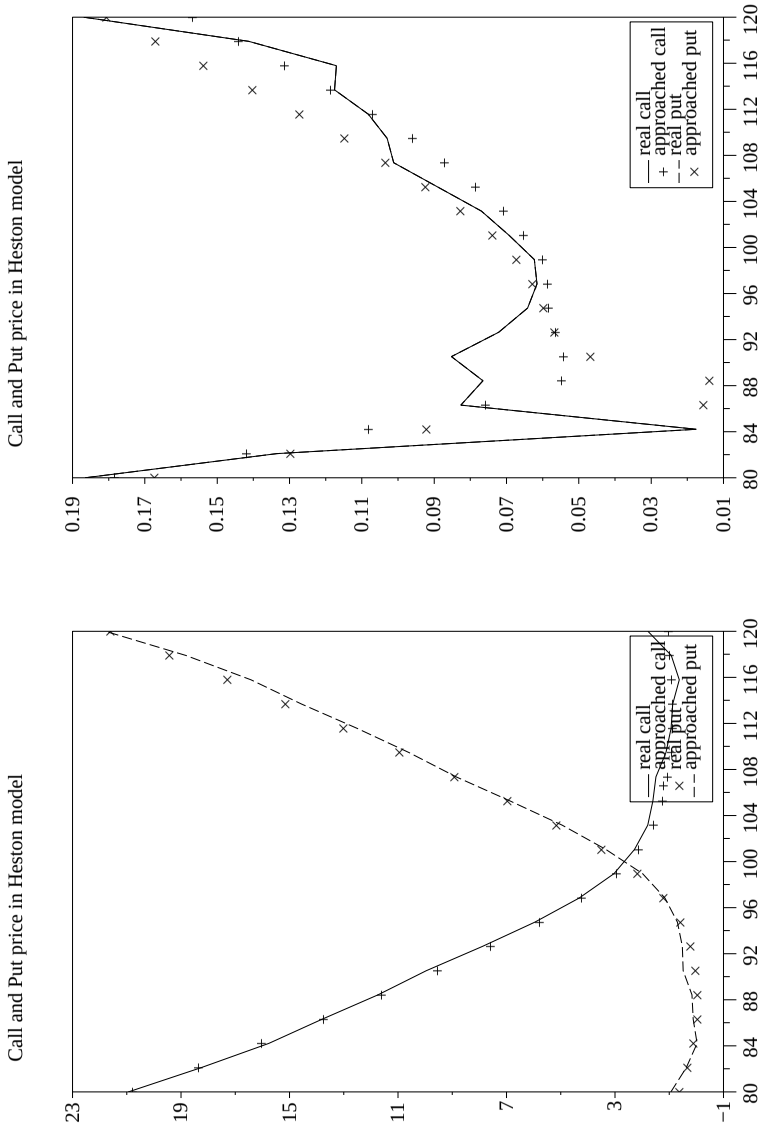

Figure 12: Computation in "Heston" model: time step $= 12$, 20 experimental datas, $\sigma = 1$.

References

[1] L. Andersen and J. Andreasen. Jump diffusion models: volatility smile fitting and numerical method for pricing. *Review of Derivatives Research*, 2000.

[2] M. Avellaneda, C. Friedman, R. Holmes, and D. Samperi. Calibrating volatility surfaces via relative entropy minimization. *Applied Mathematical Finance*, 4, 1997.

[3] R.M. Blumenthal and R.K. Getoor. *Markov processes and potential theory.* Academic press, New York, 1968.

[4] T.F. Coleman, Y. Li, and A. Verma. Reconstructing the unknown local volatility function. *Journal of Computational Finance*, 2:77–102, 1999.

[5] R. Cont and P. Tankov. Calibration of jump-diffusion option-pricing models: a robust non-parametric approach. Technical report, Ecole Polytechnique, 2002.

[6] E. Derman and I. Kani. Riding on the smile. *Risk, 7*, pages 32–39, 1994.

[7] B. Dupire. Pricing with a smile. *Risk*, pages 18–20, January 1994.

[8] B. Dupire. Pricing and hedging with smiles. *Mathematics of derivatives securities*, pages 103–111, 1997.

[9] M. Fuckushima, Y. Oshima, and M. Takeda. *Dirichlet forms and symmetic Markov processes.* Walter de gruyer, Berlin-New York, 1994.

[10] N. Jackson, E. Süli, and S. Howison. Computation of deterministic volatility surfaces. *Journal of Computational Finance*, 2:5–32, 1998.

[11] R. Lagnado and S. Osher. A technique for calibrating derivative security pricing models: numerical solution of an inverse problem. *J. Comp. Fin.*, 1:13–25, 1997.

[12] D. Samperi. Calibrating a diffusion pricing model with uncertain volatility: Regularization and stability. *Mathematical Finance*, 12(1):71–87, 2002.

[13] M. Sharp. *General Theory of Markov Processes.* Academic Press, Boston - San Diego - New York, 1988.

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399